



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>7</sup> : <b>C12N 15/54, 15/55, 9/12, 9/16, C07K 16/18, 16/40, C12Q 1/68, A61K 38/45, 38/46</b>		<b>A2</b>	(11) International Publication Number: <b>WO 00/55332</b>
			(43) International Publication Date: 21 September 2000 (21.09.00)
(21) International Application Number: <b>PCT/US00/07277</b>		(72) Inventors; and	
(22) International Filing Date: 17 March 2000 (17.03.00)		(75) Inventors/Applicants (for US only): <b>BANDMAN, Olga [US/US]; 366 Anna Avenue, Mountain View, CA 94043 (US). TANG, Y., Tom [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). HILLMAN, Jennifer, L. [US/US]; 230 Monroe Drive, #12, Mountain View, CA 94040 (US). BAUGHN, Mariah, R. [US/US]; 14244 Santiago Road, Sunnyvale, CA 94577 (US). AZIMZAI, Yalda [US/US]; 2045 Rock Springs Drive, Hayward, CA 94545 (US). LU, Dyung, Aina, M. [US/US]; 55 Park Belmont Place, San Jose, CA 95136 (US). AU-YOUNG, Janice [US/US]; 233 Golden Eagle Lane, Brisbane, CA 94005 (US).</b>	
(30) Priority Data:			
60/125,593	18 March 1999 (18.03.99)	US	
60/135,049	20 May 1999 (20.05.99)	US	
60/143,188	9 July 1999 (09.07.99)	US	
(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications			
US	60/135,049 (CIP)		
Filed on	20 May 1999 (20.05.99)		
US	60/143,188 (CIP)		
Filed on	9 July 1999 (09.07.99)		
US	60/125,593 (CIP)		
Filed on	18 March 1999 (18.03.99)		
(71) Applicant (for all designated States except US): <b>INCYTE PHARMACEUTICALS, INC. [US/US]; 3160 Porter Drive, Palo Alto, CA 94304 (US).</b>		(81) Designated States: <b>AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).</b>	
		<b>Published</b> <i>Without international search report and to be republished upon receipt of that report.</i>	
(54) Title: <b>REGULATORS OF INTRACELLULAR PHOSPHORYLATION</b>			
(57) Abstract			
<p>The invention provides human regulators of intracellular phosphorylation (HRIP) and polynucleotides which identify and encode HRIP. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with expression of HRIP.</p>			

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

## REGULATORS OF INTRACELLULAR PHOSPHORYLATION

### TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of regulators of intracellular phosphorylation and to the use of these sequences in the diagnosis, treatment, and prevention of neurological, cell proliferative, and autoimmune/inflammatory disorders.

### BACKGROUND OF THE INVENTION

Reversible protein phosphorylation is the main strategy for controlling the activities of eukaryotic cells. Kinases and phosphatases regulate reversible phosphorylation reactions, and are thus critical components of intracellular signal transduction pathways. Protein kinases transfer the high energy phosphate from adenosine triphosphate (ATP) to specific protein targets in response to extracellular signals (such as hormones, neurotransmitters, and growth and differentiation factors), cell cycle checkpoints (for example, signals associated with mitotic events), and environmental or nutritional stresses. Protein phosphatases mediate kinase effects by removing phosphate groups from proteins.

It is estimated that more than 1000 of the 10,000 proteins active in a typical mammalian cell are phosphorylated. In general, protein activity is stimulated by phosphorylation, and this is analogous to turning on a molecular switch. When the switch is turned on, the appropriate protein kinase activates a metabolic enzyme, regulatory protein, receptor, cytoskeletal protein, ion channel or pump, or transcription factor. Protein activity is repressed by dephosphorylation when down-regulation of a signaling pathway is required. The coordinate activities of kinases and phosphatases regulate key processes such as cell proliferation, cell differentiation, cell-cell communication, and cell cycle progression. Uncontrolled signaling has been implicated in a variety of disease conditions including inflammation, cancer, arteriosclerosis, and psoriasis.

Protein kinases phosphorylate protein acceptor molecules on hydroxylated amino acids. These kinases comprise the largest known protein group, a superfamily of enzymes with widely varied functions and specificities. Protein kinases are usually named after substrates, regulatory molecules, or some aspect of a mutant phenotype. With regard to substrates, protein kinases may be roughly divided into two groups: those that phosphorylate tyrosine residues (protein tyrosine kinases (PTKs)), and those that phosphorylate serine or threonine residues (serine/threonine kinases (STKs)). A few protein kinases have dual specificity and phosphorylate serine, threonine, and tyrosine residues. Some STKs and PTKs possess structural characteristics of both families (Hardie, G. and S. Hanks (1995) The Protein Kinase Facts Book, Vol. I:7-20, Academic Press, San Diego CA).

Almost all kinases contain a conserved 250-300 amino acid kinase domain that folds into a

two-lobed structure. The primary structure of the kinase domain is conserved and can be further subdivided into 11 subdomains. The smaller N-terminal lobe of the kinase domain, which contains subdomains I through IV, is primarily involved in the binding and orientation of the ATP (or GTP) donor molecule. The larger C terminal lobe, which contains subdomains VI through XI, binds the protein substrate and carries out transfer of the gamma phosphate from ATP to the hydroxyl group of a serine, threonine, or tyrosine residue. Subdomain V spans the two lobes. Each of the 11 subdomains contains specific amino acid residues and motifs that are characteristic of the particular subdomain and are highly conserved (Hardie, G. and S. Hanks (1995) The Protein Kinase Facts Book, Vol. I:7-20, Academic Press, San Diego CA). In particular, two protein kinase signature sequences have been identified in the kinase domain, the first containing an active site lysine residue involved in ATP binding (subdomain II), and the second containing an aspartate residue important for catalytic activity (subdomain VI). Kinases may also be categorized into families by the different amino acid sequences (generally between 5 and 100 residues) located on either side of, or inserted into, the kinase domain. These additional amino acid sequences are involved in the regulation of each kinase as the kinase recognizes and interacts with its target protein.

PTKs may be classified as either transmembrane or non-transmembrane proteins. Transmembrane PTKs function as receptors for most growth factors. Growth factors bind to the receptor tyrosine kinase (RTK), causing the RTK to phosphorylate itself (autophosphorylation) and specific intracellular second messenger proteins. Growth factors that bind RTKs include epidermal growth factor, platelet-derived growth factor, fibroblast growth factor, hepatocyte growth factor, insulin and insulin-like growth factors, nerve growth factor, vascular endothelial growth factor, and macrophage colony stimulating factor.

Non-transmembrane PTKs form signaling complexes with the cytosolic domains of plasma membrane receptors. Receptors that signal through non-transmembrane PTKs include receptors for cytokines and hormones (e.g., growth hormone and prolactin), and antigen-specific receptors on T and B lymphocytes. Many non-transmembrane PTKs were first identified as the products of mutant oncogenes in cancer cells in which PTK activation was no longer subject to normal cellular controls. About one third of the known oncogenes encode PTKs, and it is well known that cellular transformation (oncogenesis) is often accompanied by increased tyrosine phosphorylation activity (Charbonneau H and N.K. Tonks (1992) *Annu. Rev. Cell Biol.* 8:463-493). Regulation of PTK activity may therefore be an important strategy in controlling some types of cancer.

STKs are non-transmembrane proteins. STKs include second messenger-dependent protein kinases, which primarily mediate the effects of second messengers such as cyclic AMP, cyclic GMP, inositol triphosphate, phosphatidylinositol 3,4,5-triphosphate, cyclic ADP ribose, arachidonic acid, diacylglycerol, and calcium-calmodulin (CaM). STKs include cyclic AMP dependent protein kinases

(PKAs), which are involved in mediating hormone-induced cellular responses; CaM-dependent protein kinases, which are involved in regulation of smooth muscle contraction, glycogen breakdown, and neurotransmission; mitogen-activated protein (MAP) kinases, which mediate signal transduction from the cell surface to the nucleus via phosphorylation cascades; and diacylglycerol-activated protein kinase C (PKC), which mediates glycogen breakdown and activation of various transcription factors. PKC mu is a novel member of the PKC family that, like other PKCs, contains a zinc-finger-like, cysteine-rich motif in the N-terminal region necessary for phorbol ester binding, and is also capable of phorbol ester-independent kinase activity (Johannes, F.J. et al. (1994) J. Biol. Chem. 269:6140-6148).

10           The PKAs are activated by cAMP produced within the cell in response to hormone stimulation. Altered PKA expression is implicated in a variety of disorders and diseases including cancer, thyroid disorders, diabetes, atherosclerosis, and cardiovascular disease (Isselbacher, K.J. et al. (1994) Harrison's Principles of Internal Medicine, McGraw-Hill, New York NY, pp. 416-431, 1987). CaM dependent protein kinases are activated by calmodulin, an intracellular calcium receptor, in response to the concentration of free calcium in the cell. CaM kinase I and CaM kinase II play important roles in the regulation of neurotransmission, and kinases have been associated with neurological disorders such as Alzheimer's disease and with cognitive effects of aging. (See, for example, Lynch, M.A. (1998) Prog. Neurobiol. 56:571-589 and Bonkale, W.L. et al. (1999) Brain Res. 818:383-396.) CASK is a neuronal cell surface protein (neurexin) that includes a calmodulin-dependent protein kinase domain and is present in relatively high concentrations in brain synaptic plasma membranes (Hata, Y. et al. (1996) J. Neurosci. 16:2488-2494). CASK forms part of a complex capable of binding the amyloid precursor protein (APP) implicated in Alzheimer's Disease, and may play an important role in APP processing (Borg, J.P. et al. (1998) J. Biol. Chem. 273:31633-31636).

25           The cyclin-dependent protein kinases (CDKs) are STKs that control the progression of cells through the cell cycle. Cyclins are small regulatory proteins that bind to and activate CDKs, which then phosphorylate and activate selected proteins involved in the mitotic process. CDKs are unique in that they require multiple inputs to become activated. In addition to cyclin, CDK activation requires the phosphorylation of a specific threonine residue and the dephosphorylation of a specific tyrosine residue. Another family of STKs associated with the cell cycle are the NIMA (never in mitosis)-related kinases (Neks). Both CDKs and Neks are involved in duplication, maturation, and separation of the microtubule organizing center, the centrosome, in animal cells (Fry, A.M. et al. (1998) EMBO J. 17:470-481).

35           The MAP kinases comprise another STK family that regulates intracellular signaling pathways. The MAP kinases mediate signal transduction from the cell surface to the nucleus via

phosphorylation cascades. The extracellular stimuli which activate MAP kinase pathways include epidermal growth factor, ultraviolet light, hyperosmolar medium, heat shock, endotoxic lipopolysaccharide, and pro-inflammatory cytokines such as tumor necrosis factor and interleukin-1. Altered MAP kinase expression is implicated in a variety of disease conditions including cancer, inflammation, immune disorders, and disorders affecting growth and development.

Apoptosis is a highly regulated signaling pathway leading to cell death that plays a crucial role in tissue development and homeostasis. Deregulation of this process is associated with the pathogenesis of a number of diseases including autoimmune disease, neurodegenerative disorders, and cancer. Various STKs play key roles in this process. ZIP kinase is an STK containing a C-terminal leucine zipper domain in addition to its N-terminal protein kinase domain. This C-terminal domain appears to mediate homodimerization and activation of the kinase as well as interactions with transcription factors such as activating transcription factor, ATF4, a member of the cyclic-AMP responsive element binding protein (ATF/CREB) family of transcriptional factors (Sanjo, H. et al. (1998) J. Biol. Chem, 273:29066-29071). DRAK1 and DRAK2 are STKs that share homology with the death-associated protein kinases (DAP kinases), known to function in interferon- $\gamma$  induced apoptosis (Sanjo et al., supra). Like ZIP kinase, DAP kinases contain a C-terminal protein-protein interaction domain, in the form of ankyrin repeats, in addition to the N-terminal kinase domain. These types of kinases, ZIP, DAP, and DRAK, induce morphological changes associated with apoptosis when transfected into NIH3T3 cells (Sanjo et al., supra). However, deletion of either the N-terminal kinase catalytic domain or the C-terminal domain of these proteins abolishes apoptotic activity, indicating that in addition to the kinase activity, activity in the C-terminal domain is also necessary for apoptosis, possibly as an interacting domain with a regulator or a specific substrate.

RICK is another STK recently identified as mediating a specific apoptotic pathway involving the death receptor, CD95 (Inohara, N. et al. (1998) J. Biol. Chem. 273:12296-12300). CD95 is a member of the tumor necrosis factor receptor superfamily and plays a critical role in the regulation and homeostasis of the immune system (Nagata, S. (1997) Cell 88:355-365). The CD95 receptor signaling pathway involves recruitment of various intracellular molecules to a receptor complex following ligand binding. This includes recruitment of the cysteine protease caspase-8 which, in turn, activates a caspase cascade leading to cell death. RICK is composed of an N-terminal kinase catalytic domain and a C-terminal "caspase-recruitment" domain that interacts with caspase-like domains, indicating that RICK plays a role in the recruitment of caspase-8. This interpretation is supported by the fact that the expression of RICK in human 293T cells promotes activation of caspase-8 and potentiates the induction of apoptosis by various proteins involved in the CD95 apoptotic pathway (Inohara et al., supra).

Protein phosphatases regulate the effects of protein kinases by removing phosphate groups from molecules activated by kinases. Phosphatases are characterized as either tyrosine-specific or serine/threonine-specific based on their preferred phospho-amino acid substrate, although some protein phosphatases have dual specificity for both serine/threonine and tyrosine groups.

- 5 Serine/threonine phosphatases dephosphorylate phosphoserine and phosphothreonine residues, and are important regulators of many cAMP-mediated hormone responses (Cohen, P. (1989) *Annu. Rev. Biochem.* 58:453-508). Serine/threonine phosphatases generally comprise two or more subunits and have broad and overlapping protein substrate specificities.

- 10 Tyrosine phosphatases are generally monomeric proteins which function primarily in the transduction of signals across the plasma membrane, and are categorized as either transmembrane receptor-like proteins or soluble non-transmembrane proteins. Tyrosine phosphatases reverse the effects of PTKs, removing phosphate groups from tyrosine residues of phosphorylated proteins, and play a significant role in cell cycle and cell signaling processes, lymphocyte activation, and cell adhesion (Charbonneau, H. and N.K. Tonks (1992) *Annu. Rev. Cell Biol.* 8:463-493). In the process  
15 of cell division, for example, a specific tyrosine phosphatase called M-phase inducer phosphatase plays a key role in the induction of mitosis by dephosphorylating and activating CDC2, a cell division-specific PTK (Krishna, S. et al. (1990) *Proc. Natl. Acad. Sci. USA* 87:5139-5143).

- Tyrosine phosphatases share a conserved catalytic domain of about 250 amino acids which contains the active site. The active site consensus sequence consists of 13 amino acids, including a  
20 cysteine residue that is essential for phosphatase activity. In addition, the genes encoding at least eight tyrosine phosphatases have been mapped to chromosomal regions that are translocated or rearranged in various neoplastic conditions, including lymphoma, leukemia, small cell lung carcinoma, adenocarcinoma, and neuroblastoma (Charbonneau, H. and N.K. Tonks (1992) *Annu. Rev. Cell Biol.* 8:463-493). As previously noted, many PTKs are encoded by oncogenes, and  
25 oncogenesis is often accompanied by increased tyrosine phosphorylation activity. It is therefore possible that tyrosine phosphatases may prevent or reverse cell transformation and the growth of various cancers by controlling the levels of tyrosine phosphorylation in cells. This hypothesis is supported by studies showing that overexpression of tyrosine phosphatases can suppress transformation in cells, and that specific inhibition of tyrosine phosphatases can enhance cell  
30 transformation (Charbonneau and Tonks, *supra*).

- In addition to protein phosphorylation, lipid phosphorylation also plays a role in certain signaling pathways. The phosphorylation of phosphatidylinositol is involved in activation of the PKC signaling pathway. Recently, a sphingolipid metabolite, sphingosine-1-phosphate (SPP), has emerged as a novel lipid second-messenger with both extracellular and intracellular actions (Kohama,  
35 T. et al. (1998) *J. Biol. Chem.* 273:23722-23728). Extracellularly, SPP is a ligand for the G-protein

coupled receptor EDG-1 (endothelial-derived, G-protein coupled receptor). Intracellularly, SPP regulates cell growth, survival, motility, and cytoskeletal changes. SPP levels are regulated by sphingosine kinases that specifically phosphorylate D-erythro-sphingosine to SPP. The importance of sphingosine kinase in cell signaling is indicated by the fact that various stimuli, including  
5 platelet-derived growth factor (PDGF), nerve growth factor, and activation of PKC, increase cellular levels of SPP by activation of sphingosine kinase, and the fact that competitive inhibitors of the enzyme selectively inhibit cell proliferation induced by PDGF (Kohama et al., supra).

The discovery of new regulators of intracellular phosphorylation and the polynucleotides encoding them satisfies a need in the art by providing new compositions which are useful in the  
10 diagnosis, prevention, and treatment of neurological, cell proliferative, and autoimmune/inflammatory disorders.

### SUMMARY OF THE INVENTION

The invention features purified polypeptides, regulators of intracellular phosphorylation,  
15 referred to collectively as "HRIP" and individually as "HRIP-1," "HRIP-2," "HRIP-3," "HRIP-4," "HRIP-5," "HRIP-6," "HRIP-7," "HRIP-8," "HRIP-9," "HRIP-10," "HRIP-11," "HRIP-12," "HRIP-13," and "HRIP-14." In one aspect, the invention provides an isolated polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-14, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from  
20 the group consisting of SEQ ID NO:1-14, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-14, or d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-14. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-14.

The invention further provides an isolated polynucleotide encoding a polypeptide comprising  
25 a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-14, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-14, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-14, or d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-14. In one  
30 alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:15-28.

Additionally, the invention provides a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding a polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-14, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group  
35 consisting of SEQ ID NO:1-14, c) a biologically active fragment of an amino acid sequence selected



from the group consisting of SEQ ID NO:1-14, or d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-14. In one alternative, the invention provides a cell transformed with the recombinant polynucleotide. In another alternative, the invention provides a transgenic organism comprising the recombinant polynucleotide.

5           The invention also provides a method for producing a polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-14, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-14, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-14, or d) an immunogenic fragment of an amino  
10 acid sequence selected from the group consisting of SEQ ID NO:1-14. The method comprises a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding the polypeptide, and b) recovering the polypeptide so expressed.

          Additionally, the invention provides an isolated antibody which specifically binds to a  
15 polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-14, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-14, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-14, or d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID  
20 NO:1-14.

          The invention further provides an isolated polynucleotide comprising a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:15-28, b) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:15-28, c) a polynucleotide sequence complementary to a),  
25 or d) a polynucleotide sequence complementary to b). In one alternative, the polynucleotide comprises at least 60 contiguous nucleotides.

          Additionally, the invention provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide comprising a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:15-28, b) a naturally  
30 occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:15-28, c) a polynucleotide sequence complementary to a), or d) a polynucleotide sequence complementary to b). The method comprises a) hybridizing the sample with a probe comprising at least 16 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically  
35 hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed

between said probe and said target polynucleotide, and b) detecting the presence or absence of said hybridization complex, and optionally, if present, the amount thereof. In one alternative, the probe comprises at least 30 contiguous nucleotides. In another alternative, the probe comprises at least 60 contiguous nucleotides.

5           The invention further provides a pharmaceutical composition comprising an effective amount of a polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-14, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-14, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-14, or d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID  
10 NO:1-14, and a pharmaceutically acceptable excipient. The invention additionally provides a method of treating a disease or condition associated with decreased expression of functional HRIP, comprising administering to a patient in need of such treatment the pharmaceutical composition.

          The invention also provides a method for screening a compound for effectiveness as an  
15 agonist of a polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-14, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-14, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-14, or d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ  
20 ID NO:1-14. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting agonist activity in the sample. In one alternative, the invention provides a pharmaceutical composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with decreased expression of functional HRIP, comprising  
25 administering to a patient in need of such treatment the pharmaceutical composition.

          Additionally, the invention provides a method for screening a compound for effectiveness as an antagonist of a polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-14, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-14,  
30 c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-14, or d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-14. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting antagonist activity in the sample. In one alternative, the invention provides a pharmaceutical composition comprising an antagonist compound identified by  
35 the method and a pharmaceutically acceptable excipient. In another alternative, the invention

provides a method of treating a disease or condition associated with overexpression of functional HRIP, comprising administering to a patient in need of such treatment the pharmaceutical composition.

5 The invention further provides a method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence selected from the group consisting of SEQ ID NO:15-28, the method comprising a) exposing a sample comprising the target polynucleotide to a compound, and b) detecting altered expression of the target polynucleotide.

## 10 BRIEF DESCRIPTION OF THE TABLES

Table 1 shows polypeptide and nucleotide sequence identification numbers (SEQ ID NOs), clone identification numbers (clone IDs), cDNA libraries, and cDNA fragments used to assemble full-length sequences encoding HRIP.

15 Table 2 shows features of each polypeptide sequence, including potential motifs, homologous sequences, and methods, algorithms, and searchable databases used for analysis of HRIP.

Table 3 shows selected fragments of each nucleic acid sequence; the tissue-specific expression patterns of each nucleic acid sequence as determined by northern analysis; diseases, disorders, or conditions associated with these tissues; and the vector into which each cDNA was cloned.

20 Table 4 describes the tissues used to construct the cDNA libraries from which cDNA clones encoding HRIP were isolated.

Table 5 shows the tools, programs, and algorithms used to analyze HRIP, along with applicable descriptions, references, and threshold parameters.

## 25 DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which  
30 will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so  
35 forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

## DEFINITIONS

“HRIP” refers to the amino acid sequences of substantially purified HRIP obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and human, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term “agonist” refers to a molecule which intensifies or mimics the biological activity of HRIP. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of HRIP either by directly interacting with HRIP or by acting on components of the biological pathway in which HRIP participates.

An “allelic variant” is an alternative form of the gene encoding HRIP. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. A gene may have none, one, or many allelic variants of its naturally occurring form. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

“Altered” nucleic acid sequences encoding HRIP include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as HRIP or a polypeptide with at least one functional characteristic of HRIP. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding HRIP, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding HRIP. The encoded protein may also be “altered,” and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent HRIP. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of HRIP is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged

amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine. Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

5       The terms "amino acid" and "amino acid sequence" refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. Where "amino acid sequence" is recited to refer to an amino acid sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

10       "Amplification" relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

15       The term "antagonist" refers to a molecule which inhibits or attenuates the biological activity of HRIP. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of HRIP either by directly interacting with HRIP or by acting on components of the biological pathway in which HRIP participates.

20       The term "antibody" refers to intact immunoglobulin molecules as well as to fragments thereof, such as Fab, F(ab')<sub>2</sub>, and Fv fragments, which are capable of binding an epitopic determinant. Antibodies that bind HRIP polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

25       The term "antigenic determinant" refers to that region of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (particular regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

30       The term "antisense" refers to any composition capable of base-pairing with the "sense" strand of a specific nucleic acid sequence. Antisense compositions may include DNA; RNA; peptide nucleic acid (PNA); oligonucleotides having modified backbone linkages such as phosphorothioates,

methyolphosphonates, or benzylphosphonates; oligonucleotides having modified sugar groups such as 2'-methoxyethyl sugars or 2'-methoxyethoxy sugars; or oligonucleotides having modified bases such as 5-methyl cytosine, 2'-deoxyuracil, or 7-deaza-2'-deoxyguanosine. Antisense molecules may be produced by any method including chemical synthesis or transcription. Once introduced into a cell, the complementary antisense molecule base-pairs with a naturally occurring nucleic acid sequence produced by the cell to form duplexes which block either transcription or translation. The designation "negative" or "minus" can refer to the antisense strand, and the designation "positive" or "plus" can refer to the sense strand of a reference DNA molecule.

The term "biologically active" refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" refers to the capability of the natural, recombinant, or synthetic HRIP, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The terms "complementary" and "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence "5' A-G-T 3'" bonds to the complementary sequence "3' T-C-A 5'." Complementarity between two single-stranded molecules may be "partial," such that only some of the nucleic acids bind, or it may be "complete," such that total complementarity exists between the single stranded molecules. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands. This is of particular importance in amplification reactions, which depend upon binding between nucleic acid strands, and in the design and use of peptide nucleic acid (PNA) molecules.

A "composition comprising a given polynucleotide sequence" and a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding HRIP or fragments of HRIP may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been resequenced to resolve uncalled bases, extended using the XL-PCR kit (Perkin-Elmer, Norwalk CT) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from the overlapping sequences of one or more Incyte Clones and, in some cases, one or more public domain ESTs, using a computer program for fragment assembly, such as the GELVIEW fragment assembly system (GCG, Madison WI). Some sequences have been both extended and assembled to produce the consensus sequence.

"Conservative amino acid substitutions" are those substitutions that, when made, least interfere with the properties of the original protein, i.e., the structure and especially the function of the protein is conserved and not significantly changed by such substitutions. The table below shows amino acids which may be substituted for an original amino acid in a protein and which are regarded as conservative amino acid substitutions.

	Original Residue	Conservative Substitution
	Ala	Gly, Ser
	Arg	His, Lys
	Asn	Asp, Gln, His
10	Asp	Asn, Glu
	Cys	Ala, Ser
	Gln	Asn, Glu, His
	Glu	Asp, Gln, His
	Gly	Ala
15	His	Asn, Arg, Gln, Glu
	Ile	Leu, Val
	Leu	Ile, Val
	Lys	Arg, Gln, Glu
	Met	Leu, Ile
20	Phe	His, Met, Leu, Trp, Tyr
	Ser	Cys, Thr
	Thr	Ser, Val
	Trp	Phe, Tyr
	Tyr	His, Phe, Trp
25	Val	Ile, Leu, Thr

Conservative amino acid substitutions generally maintain (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a beta sheet or alpha helical conformation, (b) the charge or hydrophobicity of the molecule at the site of the substitution, and/or (c) the bulk of the side chain.

A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to the chemical modification of a polypeptide sequence, or a polynucleotide sequence. Chemical modifications of a polynucleotide sequence can include, for example, replacement of hydrogen by an alkyl, acyl, hydroxyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

A "fragment" is a unique portion of HRIP or the polynucleotide encoding HRIP which is identical in sequence to but shorter in length than the parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example,

a fragment may comprise from 5 to 1000 contiguous nucleotides or amino acid residues. A fragment used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid residues in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50% of a polypeptide) as shown in a certain defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the present embodiments.

10 A fragment of SEQ ID NO:15-28 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:15-28, for example, as distinct from any other sequence in the same genome. A fragment of SEQ ID NO:15-28 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:15-28 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:15-28 and the region of  
15 SEQ ID NO:15-28 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-14 is encoded by a fragment of SEQ ID NO:15-28. A fragment of SEQ ID NO:1-14 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-14. For example, a fragment of SEQ ID NO:1-14 is useful as an immunogenic peptide  
20 for the development of antibodies that specifically recognize SEQ ID NO:1-14. The precise length of a fragment of SEQ ID NO:1-14 and the region of SEQ ID NO:1-14 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

The term "similarity" refers to a degree of complementarity. There may be partial similarity  
25 or complete similarity. The word "identity" may substitute for the word "similarity." A partially complementary sequence that at least partially inhibits an identical sequence from hybridizing to a target nucleic acid is referred to as "substantially similar." The inhibition of hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization assay (Southern or northern blot, solution hybridization, and the like) under conditions of reduced  
30 stringency. A substantially similar sequence or hybridization probe will compete for and inhibit the binding of a completely similar (identical) sequence to the target sequence under conditions of reduced stringency. This is not to say that conditions of reduced stringency are such that non-specific binding is permitted, as reduced stringency conditions require that the binding of two sequences to one another be a specific (i.e., a selective) interaction. The absence of non-specific binding may be  
35 tested by the use of a second target sequence which lacks even a partial degree of complementarity



(e.g., less than about 30% similarity or identity). In the absence of non-specific binding, the substantially similar sequence or probe will not hybridize to the second non-complementary target sequence.

5 The phrases "percent identity" and "% identity," as applied to polynucleotide sequences, refer to the percentage of residue matches between at least two polynucleotide sequences aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps in the sequences being compared in order to optimize alignment between two sequences, and therefore achieve a more meaningful comparison of the two sequences.

10 Percent identity between polynucleotide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS 8:189-191. For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polynucleotide sequence pairs.

Alternatively, a suite of commonly used and freely available sequence comparison algorithms is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at <http://www.ncbi.nlm.nih.gov/BLAST/>. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at <http://www.ncbi.nlm.nih.gov/gorf/bl2.html>. The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such default parameters may be, for example:

*Matrix: BLOSUM62*

*Reward for match: 1*

*Penalty for mismatch: -2*

*Open Gap: 5 and Extension Gap: 2 penalties*

35 *Gap x drop-off: 50*

*Expect: 10*

*Word Size: 11*

*Filter: on*

Percent identity may be measured over the length of an entire defined sequence, for example,  
5 as defined by a particular SEQ ID number, or may be measured over a shorter length, for example,  
over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at  
least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous  
nucleotides. Such lengths are exemplary only, and it is understood that any fragment length  
supported by the sequences shown herein, in the tables, figures, or Sequence Listing, may be used to  
10 describe a length over which percentage identity may be measured.

Nucleic acid sequences that do not show a high degree of identity may nevertheless encode  
similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes  
in a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid  
sequences that all encode substantially the same protein.

15 The phrases "percent identity" and "% identity," as applied to polypeptide sequences, refer to  
the percentage of residue matches between at least two polypeptide sequences aligned using a  
standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some  
alignment methods take into account conservative amino acid substitutions. Such conservative  
substitutions, explained in more detail above, generally preserve the hydrophobicity and acidity at the  
20 site of substitution, thus preserving the structure (and therefore function) of the polypeptide.

Percent identity between polypeptide sequences may be determined using the default  
parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e  
sequence alignment program (described and referenced above). For pairwise alignments of  
polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap  
25 penalty=3, window=5, and "diagonals saved"=5. The PAM250 matrix is selected as the default  
residue weight table. As with polynucleotide alignments, the percent identity is reported by  
CLUSTAL V as the "percent similarity" between aligned polypeptide sequence pairs.

Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise  
comparison of two polypeptide sequences, one may use the "BLAST 2 Sequences" tool Version 2.0.9  
30 (May-07-1999) with blastp set at default parameters. Such default parameters may be, for example:

*Matrix: BLOSUM62*

*Open Gap: 11 and Extension Gap: 1 penalties*

*Gap x drop-off: 50*

*Expect: 10*

35 *Word Size: 3*

*Filter: on*

Percent identity may be measured over the length of an entire defined polypeptide sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequence, for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

“Human artificial chromosomes” (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size, and which contain all of the elements required for stable mitotic chromosome segregation and maintenance.

The term “humanized antibody” refers to antibody molecules in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

“Hybridization” refers to the process by which a polynucleotide strand anneals with a complementary strand through base pairing under defined hybridization conditions. Specific hybridization is an indication that two nucleic acid sequences share a high degree of identity. Specific hybridization complexes form under permissive annealing conditions and remain hybridized after the “washing” step(s). The washing step(s) is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 µg/ml denatured salmon sperm DNA.

Generally, stringency of hybridization is expressed, in part, with reference to the temperature under which the wash step is carried out. Generally, such wash temperatures are selected to be about 5°C to 20°C lower than the thermal melting point ( $T_m$ ) for the specific sequence at a defined ionic strength and pH. The  $T_m$  is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. An equation for calculating  $T_m$  and conditions for nucleic acid hybridization are well known and can be found in Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, 2<sup>nd</sup> ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; specifically see volume 2, chapter 9.

High stringency conditions for hybridization between polynucleotides of the present

invention include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS, for 1 hour. Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC concentration may be varied from about 0.1 to 2 x SSC, with SDS being present at about 0.1%. Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance, denatured salmon sperm DNA at about 100-200 µg/ml. Organic solvent, such as formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, such as for RNA:DNA hybridizations. Useful variations on these wash conditions will be readily apparent to those of ordinary skill in the art. Hybridization, particularly under high stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C<sub>0</sub>t or R<sub>0</sub>t analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

The words "insertion" and "addition" refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively.

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

An "immunogenic fragment" is a polypeptide or oligopeptide fragment of HRIP which is capable of eliciting an immune response when introduced into a living organism, for example, a mammal. The term "immunogenic fragment" also includes any polypeptide or oligopeptide fragment of HRIP which is useful in any of the antibody production methods disclosed herein or known in the art.

The term "microarray" refers to an arrangement of distinct polynucleotides on a substrate.

The terms "element" and "array element" in a microarray context, refer to hybridizable polynucleotides arranged on the surface of a substrate.

The term "modulate" refers to a change in the activity of HRIP. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of HRIP.

The phrases "nucleic acid" and "nucleic acid sequence" refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or

synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

"Operably linked" refers to the situation in which a first nucleic acid sequence is placed in a functional relationship with the second nucleic acid sequence. For instance, a promoter is operably  
5 linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Generally, operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of  
10 amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

"Probe" refers to nucleic acid sequences encoding HRIP, their complements, or fragments thereof, which are used to detect identical, allelic or related nucleic acid sequences. Probes are  
15 isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes.

"Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and  
20 identification) of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR).

Probes and primers as used in the present invention typically comprise at least 15 contiguous nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also be employed, such as probes and primers that comprise at least 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences. Probes and primers  
25 may be considerably longer than these examples, and it is understood that any length supported by the specification, including the tables, figures, and Sequence Listing, may be used.

Methods for preparing and using probes and primers are described in the references, for example Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, 2<sup>nd</sup> ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; Ausubel et al., 1987, Current Protocols in Molecular Biology,  
30 Greene Publ. Assoc. & Wiley-Intersciences, New York NY; Innis et al., 1990, PCR Protocols, A Guide to Methods and Applications, Academic Press, San Diego CA. PCR primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA).

Oligonucleotides for use as primers are selected using software known in the art for such  
35 purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to

100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

A "recombinant nucleic acid" is a sequence that is not naturally occurring or has a sequence that is made by an artificial combination of two or more otherwise separated segments of sequence. This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques such as those described in Sambrook, supra. The term recombinant includes nucleic acids that have been altered solely by addition, substitution, or deletion of a portion of the nucleic acid. Frequently, a recombinant nucleic acid may include a nucleic acid sequence operably linked to a promoter sequence. Such a recombinant nucleic acid may be part of a vector that is used, for example, to transform a cell.

Alternatively, such recombinant nucleic acids may be part of a viral vector, e.g., based on a vaccinia virus, that could be used to vaccinate a mammal wherein the recombinant nucleic acid is expressed, inducing a protective immunological response in the mammal.

An "RNA equivalent," in reference to a DNA sequence, is composed of the same linear sequence of nucleotides as the reference DNA sequence with the exception that all occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose

instead of deoxyribose.

The term "sample" is used in its broadest sense. A sample suspected of containing nucleic acids encoding HRIP, or fragments thereof, or HRIP itself, may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" and "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, an antagonist, a small molecule, or any natural or synthetic binding composition. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide containing the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least 60% free, preferably at least 75% free, and most preferably at least 90% free from other components with which they are naturally associated.

A "substitution" refers to the replacement of one or more amino acids or nucleotides by different amino acids or nucleotides, respectively.

"Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

"Transformation" describes a process by which exogenous DNA enters and changes a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term "transformed" cells includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A "transgenic organism," as used herein, is any organism, including but not limited to animals and plants, in which one or more of the cells of the organism contains heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor

of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. The term genetic manipulation does not include classical cross-breeding, or in vitro fertilization, but rather is directed to the introduction of a recombinant DNA molecule. The transgenic organisms contemplated in accordance with the present invention include bacteria, cyanobacteria, fungi, and plants and animals. The isolated DNA of the present invention can be introduced into the host by methods known in the art, for example infection, transfection, transformation or transconjugation. Techniques for transferring the DNA of the present invention into such organisms are widely known and provided in references such as Sambrook et al. (1989), supra.

10 A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95% or at least 98% or  
15 greater sequence identity over a certain defined length. A variant may be described as, for example, an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternate splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the  
20 reference molecule. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides generally will have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The  
25 presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-  
30 1999) set at default parameters. Such a pair of polypeptides may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or at least 98% or greater sequence identity over a certain defined length of one of the polypeptides.

## THE INVENTION

The invention is based on the discovery of new human regulators of intracellular  
35 phosphorylation (HRIP), the polynucleotides encoding HRIP, and the use of these compositions for



the diagnosis, treatment, or prevention of neurological, cell proliferative, and autoimmune/inflammatory disorders.

Table 1 lists the Incyte clones used to assemble full length nucleotide sequences encoding HRIP. Columns 1 and 2 show the sequence identification numbers (SEQ ID NOs) of the polypeptide and nucleotide sequences, respectively. Column 3 shows the clone IDs of the Incyte clones in which nucleic acids encoding each HRIP were identified, and column 4 shows the cDNA libraries from which these clones were isolated. Column 5 shows Incyte clones and their corresponding cDNA libraries. Clones for which cDNA libraries are not indicated were derived from pooled cDNA libraries. The Incyte clones in column 5 were used to assemble the consensus nucleotide sequence of each HRIP and are useful as fragments in hybridization technologies.

The columns of Table 2 show various properties of each of the polypeptides of the invention: column 1 references the SEQ ID NO; column 2 shows the number of amino acid residues in each polypeptide; column 3 shows potential phosphorylation sites; column 4 shows potential glycosylation sites; column 5 shows the amino acid residues comprising signature sequences and motifs; column 6 shows homologous sequences as identified by BLAST analysis; and column 7 shows analytical methods and in some cases, searchable databases to which the analytical methods were applied. The methods of column 7 were used to characterize each polypeptide through sequence homology and protein motifs.

The columns of Table 3 show the tissue-specificity and diseases, disorders, or conditions associated with nucleotide sequences encoding HRIP. The first column of Table 3 lists the nucleotide SEQ ID NOs. Column 2 lists fragments of the nucleotide sequences of column 1. These fragments are useful, for example, in hybridization or amplification technologies to identify SEQ ID NO:15-28 and to distinguish between SEQ ID NO:15-28 and related polynucleotide sequences. For SEQ ID NO:15-27, the polypeptides encoded by these fragments are useful, for example, as immunogenic peptides. Column 3 lists tissue categories which express HRIP as a fraction of total tissues expressing HRIP. Column 4 lists diseases, disorders, or conditions associated with those tissues expressing HRIP as a fraction of total tissues expressing HRIP. Column 5 lists the vectors used to subclone each cDNA library.

The columns of Table 4 show descriptions of the tissues used to construct the cDNA libraries from which cDNA clones encoding HRIP were isolated. Column 1 references the nucleotide SEQ ID NOs, column 2 shows the cDNA libraries from which these clones were isolated, and column 3 shows the tissue origins and other descriptive information relevant to the cDNA libraries in column 2.

The invention also encompasses HRIP variants. A preferred HRIP variant is one which has at least about 80%, or alternatively at least about 90%, or even at least about 95% amino acid sequence identity to the HRIP amino acid sequence, and which contains at least one functional or structural

characteristic of HRIP.

The invention also encompasses polynucleotides which encode HRIP. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:15-28, which encodes HRIP. The polynucleotide sequences of SEQ ID NO:15-28, as presented in the Sequence Listing, embrace the equivalent RNA sequences, wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The invention also encompasses a variant of a polynucleotide sequence encoding HRIP. In particular, such a variant polynucleotide sequence will have at least about 80%, or alternatively at least about 90%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding HRIP. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:15-28 which has at least about 80%, or alternatively at least about 90%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:15-28. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of HRIP.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding HRIP, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring HRIP, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode HRIP and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring HRIP under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding HRIP or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding HRIP and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode HRIP and

HRIP derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding HRIP or any fragment thereof.

5 Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:15-28 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) *Methods Enzymol.* 152:399-407; Kimmel, A.R. (1987) *Methods Enzymol.* 152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in  
10 "Definitions."

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Perkin-Elmer), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or  
15 combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (Perkin-Elmer). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing  
20 system (Perkin-Elmer), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

25 The nucleic acid sequences encoding HRIP may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) *PCR Methods Applic.* 2:318-322.)  
30 Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) *Nucleic Acids Res.* 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) *PCR Methods Applic.* 1:111-119.) In this method, multiple restriction enzyme  
35

digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries  
5 (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 Primer Analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of  
10 about 68°C to 72°C.

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence  
15 into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the  
20 emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Perkin-Elmer), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

25 In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode HRIP may be cloned in recombinant DNA molecules that direct expression of HRIP, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express HRIP.

30 The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter HRIP-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-  
35 mediated site-directed mutagenesis may be used to introduce mutations that create new restriction

sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

The nucleotides of the present invention may be subjected to DNA shuffling techniques such as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent Number 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat. Biotechnol. 17:259-264; and Cramer, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or improve the biological properties of HRIP, such as its biological or enzymatic activity or its ability to bind to other molecules or compounds. DNA shuffling is a process by which a library of gene variants is produced using PCR-mediated recombination of gene fragments. The library is then subjected to selection or screening procedures that identify those gene variants with the desired properties. These preferred variants may then be pooled and further subjected to recursive rounds of DNA shuffling and selection/screening. Thus, genetic diversity is created through "artificial" breeding and rapid molecular evolution. For example, fragments of a single gene containing random point mutations may be recombined, screened, and then reshuffled until the desired properties are optimized. Alternatively, fragments of a given gene may be recombined with fragments of homologous genes in the same gene family, either from the same or different species, thereby maximizing the genetic diversity of multiple naturally occurring genes in a directed and controllable manner.

In another embodiment, sequences encoding HRIP may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) Nucleic Acids Symp. Ser. 7:215-223; and Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232.) Alternatively, HRIP itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solid-phase techniques. (See, e.g., Roberge, J.Y. et al. (1995) Science 269:202-204.) Automated synthesis may be achieved using the ABI 431A peptide synthesizer (Perkin-Elmer). Additionally, the amino acid sequence of HRIP, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g., Chiez, R.M. and F.Z. Regnier (1990) Methods Enzymol. 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY.)

In order to express a biologically active HRIP, the nucleotide sequences encoding HRIP or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and

inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding HRIP. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding HRIP. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where  
5 sequences encoding HRIP and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both  
10 natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) *Results Probl. Cell Differ.* 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding HRIP and appropriate transcriptional and translational control  
15 elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences  
20 encoding HRIP. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or  
25 animal cell systems. The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding HRIP. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding HRIP can be achieved using a multifunctional E. coli vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPT1  
30 plasmid (Life Technologies). Ligation of sequences encoding HRIP into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) *J. Biol.*  
35 *Chem.* 264:5503-5509.) When large quantities of HRIP are needed, e.g. for the production of

antibodies, vectors which direct high level expression of HRIP may be used. For example, vectors containing the strong, inducible T5 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of HRIP. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, supra; Bitter, G.A. et al. (1987) *Methods Enzymol.* 153:516-544; and Scorer, C.A. et al. (1994) *Bio/Technology* 12:181-184.)

Plant systems may also be used for expression of HRIP. Transcription of sequences encoding HRIP may be driven viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) *EMBO J.* 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) *EMBO J.* 3:1671-1680; Broglie, R. et al. (1984) *Science* 224:838-843; and Winter, J. et al. (1991) *Results Probl. Cell Differ.* 17:85-105.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding HRIP may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses HRIP in host cells. (See, e.g., Logan, J. and T. Shenk (1984) *Proc. Natl. Acad. Sci. USA* 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) *Nat. Genet.* 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of HRIP in cell lines is preferred. For example, sequences encoding HRIP can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the

introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using  
5 tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk<sup>-</sup>* and *apr<sup>-</sup>* cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic,  
10 or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which  
15 alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech),  $\beta$  glucuronidase and its substrate  $\beta$ -glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system.  
20 (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding HRIP is inserted within a marker gene sequence, transformed cells containing sequences encoding HRIP can be identified by the absence of marker gene function. Alternatively, a  
25 marker gene can be placed in tandem with a sequence encoding HRIP under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding HRIP and that express HRIP may be identified by a variety of procedures known to those of skill in the art. These  
30 procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of HRIP using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques  
35 include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and



fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on HRIP is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ.)

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding HRIP include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding HRIP, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding HRIP may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode HRIP may be designed to contain signal sequences which direct secretion of HRIP through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid

sequences encoding HRIP may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric HRIP protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of HRIP activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the HRIP encoding sequence and the heterologous protein sequence, so that HRIP may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, *supra*, ch. 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled HRIP may be achieved *in vitro* using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, for example, <sup>35</sup>S-methionine.

Fragments of HRIP may be produced not only by recombinant means, but also by direct peptide synthesis using solid-phase techniques. (See, e.g., Creighton, *supra*, pp. 55-60.) Protein synthesis may be performed by manual techniques or by automation. Automated synthesis may be achieved, for example, using the ABI 431A peptide synthesizer (Perkin-Elmer). Various fragments of HRIP may be synthesized separately and then combined to produce the full length molecule.

## THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of HRIP and regulators of intracellular phosphorylation. In addition, the expression of HRIP is closely associated with neurological tissue, with cancer and other cell proliferative disorders, and with inflammation and the immune response. Therefore, HRIP appears to play a role in neurological, cell proliferative, and autoimmune/inflammatory disorders. In the treatment of disorders associated with increased HRIP expression or activity, it is desirable to decrease the expression or activity of HRIP. In the treatment of disorders associated with decreased HRIP

expression or activity, it is desirable to increase the expression or activity of HRIP.

Therefore, in one embodiment, HRIP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HRIP. Examples of such disorders include, but are not limited to, a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease; prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome; fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis; inherited, metabolic, endocrine, and toxic myopathies; myasthenia gravis, periodic paralysis; mental disorders including mood, anxiety, and schizophrenic disorders; seasonal affective disorder (SAD); akathisia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and a cancer including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, a cancer of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; and an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's

thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative  
5 colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma.

In another embodiment, a vector capable of expressing HRIP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HRIP including, but not limited to, those described above.

10 In a further embodiment, a pharmaceutical composition comprising a substantially purified HRIP in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HRIP including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of HRIP may be  
15 administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HRIP including, but not limited to, those listed above.

In a further embodiment, an antagonist of HRIP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of HRIP. Examples of such disorders include, but are not limited to, those neurological, cell proliferative, and  
20 autoimmune/inflammatory disorders described above. In one aspect, an antibody which specifically binds HRIP may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissues which express HRIP.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding HRIP may be administered to a subject to treat or prevent a disorder associated with  
25 increased expression or activity of HRIP including, but not limited to, those described above.

In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The  
30 combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of HRIP may be produced using methods which are generally known in the art. In particular, purified HRIP may be used to produce antibodies or to screen libraries of  
35 pharmaceutical agents to identify those which specifically bind HRIP. Antibodies to HRIP may also

be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are generally preferred for therapeutic use.

5 For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with HRIP or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral  
10 gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to HRIP have an amino acid sequence consisting of at least about 5 amino acids, and generally will consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or  
15 fragments are identical to a portion of the amino acid sequence of the natural protein and contain the entire amino acid sequence of a small, naturally occurring molecule. Short stretches of HRIP amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to HRIP may be prepared using any technique which provides for the  
20 production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. USA 80:2026-2030; and Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120.)

25 In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single  
30 chain antibodies may be adapted, using methods known in the art, to produce HRIP-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137.)

Antibodies may also be produced by inducing in vivo production in the lymphocyte  
35 population or by screening immunoglobulin libraries or panels of highly specific binding reagents as

disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299.)

Antibody fragments which contain specific binding sites for HRIP may also be generated. For example, such fragments include, but are not limited to,  $F(ab)_2$  fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the  $F(ab)_2$  fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) Science 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between HRIP and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering HRIP epitopes is generally used, but a competitive binding assay may also be employed (Pound, *supra*).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for HRIP. Affinity is expressed as an association constant,  $K_a$ , which is defined as the molar concentration of HRIP-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The  $K_a$  determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple HRIP epitopes, represents the average affinity, or avidity, of the antibodies for HRIP. The  $K_a$  determined for a preparation of monoclonal antibodies, which are monospecific for a particular HRIP epitope, represents a true measure of affinity. High-affinity antibody preparations with  $K_a$  ranging from about  $10^9$  to  $10^{12}$  L/mole are preferred for use in immunoassays in which the HRIP-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with  $K_a$  ranging from about  $10^6$  to  $10^7$  L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of HRIP, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington, DC; Liddell, J.E. and Cryer, A. (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of HRIP-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and

guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, supra, and Coligan et al. supra.)

In another embodiment of the invention, the polynucleotides encoding HRIP, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, the complement of the polynucleotide encoding HRIP may be used in situations in which it would be desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary to polynucleotides encoding HRIP. Thus, complementary molecules or fragments may be used to modulate HRIP activity, or to achieve regulation of gene function. Such technology is now well known in the art, and sense or antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding HRIP.

Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. Methods which are well known to those skilled in the art can be used to construct vectors to express nucleic acid sequences complementary to the polynucleotides encoding HRIP. (See, e.g., Sambrook, supra; Ausubel, 1995, supra.)

Genes encoding HRIP can be turned off by transforming a cell or tissue with expression vectors which express high levels of a polynucleotide, or fragment thereof, encoding HRIP. Such constructs may be used to introduce untranslatable sense or antisense sequences into a cell. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled by endogenous nucleases. Transient expression may last for a month or more with a non-replicating vector, and may last even longer if appropriate replication elements are part of the vector system.

As mentioned above, modifications of gene expression can be obtained by designing complementary sequences or antisense molecules (DNA, RNA, or PNA) to the control, 5', or regulatory regions of the gene encoding HRIP. Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, may be employed. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme

molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding HRIP.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding HRIP. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nat. Biotechnol. 15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as humans, dogs, cats, cows, horses, rabbits, and monkeys.



An additional embodiment of the invention relates to the administration of a pharmaceutical or sterile composition, in conjunction with a pharmaceutically acceptable carrier, for any of the therapeutic effects discussed above. Such pharmaceutical compositions may consist of HRIP, antibodies to HRIP, and mimetics, agonists, antagonists, or inhibitors of HRIP. The compositions may be administered alone or in combination with at least one other agent, such as a stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone, or in combination with other agents, drugs, or hormones.

The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA).

Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Pharmaceutical preparations for oral use can be obtained through combining active compounds with solid excipient and processing the resultant mixture of granules (optionally, after grinding) to obtain tablets or dragee cores. Suitable auxiliaries can be added, if desired. Suitable excipients include carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, and sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums, including arabic and tragacanth; and proteins, such as gelatin and collagen. If desired, disintegrating or solubilizing agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, and alginic acid or a salt thereof, such as sodium alginate.

Dragee cores may be used in conjunction with suitable coatings, such as concentrated sugar solutions, which may also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, i.e., dosage.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with fillers or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules,  
5 the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations suitable for parenteral administration may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions may contain substances  
10 which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils, such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate, triglycerides, or liposomes. Non-lipid polycationic amino polymers may also be used for delivery. Optionally, the suspension may also contain suitable  
15 stabilizers or agents to increase the solubility of the compounds and allow for the preparation of highly concentrated solutions.

For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention may be manufactured in a manner  
20 that is known in the art, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes.

The pharmaceutical composition may be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, and succinic acids. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding  
25 free base forms. In other cases, the preparation may be a lyophilized powder which may contain any or all of the following: 1 mM to 50 mM histidine, 0.1% to 2% sucrose, and 2% to 7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. For administration of HRIP, such  
30 labeling would include amount, frequency, and method of administration.

Pharmaceutical compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

For any compound, the therapeutically effective dose can be estimated initially either in cell  
35 culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs, or pigs.

An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example HRIP or fragments thereof, antibodies of HRIP, and agonists, antagonists or inhibitors of HRIP, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the  $ED_{50}$  (the dose therapeutically effective in 50% of the population) or  $LD_{50}$  (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the  $LD_{50}/ED_{50}$  ratio. Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the  $ED_{50}$  with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about  $0.1 \mu\text{g}$  to  $100,000 \mu\text{g}$ , up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

### 30 DIAGNOSTICS

In another embodiment, antibodies which specifically bind HRIP may be used for the diagnosis of disorders characterized by expression of HRIP, or in assays to monitor patients being treated with HRIP or agonists, antagonists, or inhibitors of HRIP. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for HRIP include methods which utilize the antibody and a label to detect HRIP in human body fluids

or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

5 A variety of protocols for measuring HRIP, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of HRIP expression. Normal or standard values for HRIP expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibody to HRIP under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of HRIP expressed in subject,  
10 control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding HRIP may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect  
15 and quantify gene expression in biopsied tissues in which expression of HRIP may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of HRIP, and to monitor regulation of HRIP levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding HRIP or closely related molecules may be used to  
20 identify nucleic acid sequences which encode HRIP. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification will determine whether the probe identifies only naturally occurring sequences encoding HRIP, allelic variants, or related sequences.

25 Probes may also be used for the detection of related sequences, and may have at least 50% sequence identity to any of the HRIP encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:15-28 or from genomic sequences including promoters, enhancers, and introns of the HRIP gene.

Means for producing specific hybridization probes for DNAs encoding HRIP include the  
30 cloning of polynucleotide sequences encoding HRIP or HRIP derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as <sup>32</sup>P or <sup>35</sup>S, or by enzymatic labels,  
35 such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding HRIP may be used for the diagnosis of disorders associated with expression of HRIP. Examples of such disorders include, but are not limited to, a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other

5 extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease; prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-

10 Straussler-Scheinker syndrome; fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders,

15 peripheral nervous system disorders, dermatomyositis and polymyositis; inherited, metabolic, endocrine, and toxic myopathies; myasthenia gravis, periodic paralysis; mental disorders including mood, anxiety, and schizophrenic disorders; seasonal affective disorder (SAD); akathisia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial

20 frontotemporal dementia; a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and a cancer including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, a cancer of the adrenal gland, bladder, bone, bone marrow, brain,

25 breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; and an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis,

30 autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis,

35 myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis,

polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma. The polynucleotide sequences encoding HRIP may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered HRIP expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding HRIP may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding HRIP may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding HRIP in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of HRIP, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding HRIP, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance

of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding HRIP may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced in vitro. Oligomers will preferably contain a fragment of a polynucleotide encoding HRIP, or a fragment of a polynucleotide complementary to the polynucleotide encoding HRIP, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

Methods which may also be used to quantify the expression of HRIP include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) *J. Immunol. Methods* 159:235-244; Duplaa, C. et al. (1993) *Anal. Biochem.* 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as targets in a microarray. The microarray can be used to monitor the expression level of large numbers of genes simultaneously and to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, and to develop and monitor the activities of therapeutic agents.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) *Proc. Natl. Acad. Sci. USA* 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) *Proc. Natl. Acad. Sci. USA* 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.)

In another embodiment of the invention, nucleic acid sequences encoding HRIP may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) *Nat. Genet.* 15:345-355; Price, C.M. (1993) *Blood Rev.* 7:127-134; and Trask, B.J. (1991) *Trends Genet.* 7:149-154.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical chromosome mapping techniques and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) World Wide Web site. Correlation between the  
5 location of the gene encoding HRIP on a physical chromosomal map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder. The nucleotide sequences of the invention may be used to detect differences in gene sequences among normal, carrier, and affected individuals.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as  
10 linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the number or arm of a particular human chromosome is not known. New sequences can be assigned to chromosomal arms by physical mapping. This provides valuable information to investigators searching for disease genes using positional cloning or other  
15 gene discovery techniques. Once the disease or syndrome has been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequence of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among  
20 normal, carrier, or affected individuals.

In another embodiment of the invention, HRIP, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes  
25 between HRIP and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with HRIP, or fragments thereof,  
30 and washed. Bound HRIP is then detected by methods well known in the art. Purified HRIP can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding HRIP specifically compete with a test compound for binding HRIP. In  
35 this manner, antibodies can be used to detect the presence of any peptide which shares one or more



antigenic determinants with HRIP.

In additional embodiments, the nucleotide sequences which encode HRIP may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications, and publications mentioned above and below, in particular U.S. Ser. No. 60/125,593, U.S. Ser. No. 60/135,049, and U.S. Ser. No. 60/143,188, are hereby expressly incorporated by reference.

## EXAMPLES

### I. Construction of cDNA Libraries

RNA was purchased from Clontech or isolated from tissues described in Table 4. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A<sup>+</sup>) RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERScript plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, *supra*, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column

chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSORT1 plasmid (Life Technologies), pcDNA2.1 plasmid (Invitrogen, Carlsbad CA), or pINCY plasmid (Incyte Pharmaceuticals, Palo Alto CA). Recombinant  
5 plasmids were transformed into competent *E. coli* cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5 $\alpha$ , DH10B, or ElectroMAX DH10B from Life Technologies.

## II. Isolation of cDNA Clones

Plasmids were recovered from host cells by *in vivo* excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or  
10 WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

15 Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSKAN II fluorescence  
20 scanner (Labsystems Oy, Helsinki, Finland).

## III. Sequencing and Analysis

cDNA sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Perkin-Elmer) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific)  
25 or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the  
30 ABI PRISM 373 or 377 sequencing system (Perkin-Elmer) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997, *supra*, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example V.

35 The polynucleotide sequences derived from cDNA sequencing were assembled and analyzed

using a combination of software programs which utilize algorithms well known to those skilled in the art. Table 5 summarizes the tools, programs, and algorithms used and provides applicable descriptions, references, and threshold parameters. The first column of Table 5 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score, the greater the homology between two sequences). Sequences were analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments were generated using the default parameters specified by the clustal algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

The polynucleotide sequences were validated by removing vector, linker, and polyA sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programming, and dinucleotide nearest neighbor analysis. The sequences were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM, and PFAM to acquire annotation using programs based on BLAST, FASTA, and BLIMPS. The sequences were assembled into full length polynucleotide sequences using programs based on Phred, Phrap, and Consed, and were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length amino acid sequences, and these full length sequences were subsequently analyzed by querying against databases such as the GenBank databases (described above), SwissProt, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, and Hidden Markov Model (HMM)-based protein family databases such as PFAM. HMM is a probabilistic approach which analyzes consensus primary structures of gene families. (See, e.g., Eddy, S.R. (1996) Curr. Opin. Struct. Biol. 6:361-365.)

The programs described above for the assembly and analysis of full length polynucleotide and amino acid sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:15-28. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies were described in The Invention section above.

#### IV. Northern Analysis

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, supra, ch. 7; Ausubel,

1995, supra, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or related molecules in nucleotide databases such as GenBank or LIFESEQ (Incyte Pharmaceuticals). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$\frac{\% \text{ sequence identity} \times \% \text{ maximum BLAST score}}{100}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. For example, with a product score of 40, the match will be exact within a 1% to 2% error, and, with a product score of 70, the match will be exact. Similar molecules are usually identified by selecting those which show product scores between 15 and 40, although lower scores may identify related molecules.

The results of northern analyses are reported as a percentage distribution of libraries in which the transcript encoding HRIP occurred. Analysis involved the categorization of cDNA libraries by organ/tissue and disease. The organ/tissue categories included cardiovascular, dermatologic, developmental, endocrine, gastrointestinal, hematopoietic/immune, musculoskeletal, nervous, reproductive, and urologic. The disease/condition categories included cancer, inflammation, trauma, cell proliferation, neurological, and pooled. For each category, the number of libraries expressing the sequence of interest was counted and divided by the total number of libraries across all categories. Percentage values of tissue-specific and disease- or condition-specific expression are reported in Table 3.

#### V. Extension of HRIP Encoding Polynucleotides

The full length nucleic acid sequences of SEQ ID NO:15-28 were produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer, to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction

mix contained DNA template, 200 nmol of each primer, reaction buffer containing  $Mg^{2+}$ ,  $(NH_4)_2SO_4$ , and  $\beta$ -mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100  $\mu$ l PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5  $\mu$ l of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5  $\mu$ l to 10  $\mu$ l aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose mini-gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent *E. coli* cells. Transformed cells were selected on antibiotic-containing media, individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethylsulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer).

In like manner, the nucleotide sequences of SEQ ID NO:15-28 are used to obtain 5' regulatory sequences using the procedure above, oligonucleotides designed for such extension, and an appropriate genomic library.

#### **VI. Labeling and Use of Individual Hybridization Probes**

5 Hybridization probes derived from SEQ ID NO:15-28 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250  $\mu$ Ci of  
10  $[\gamma\text{-}^{32}\text{P}]$  adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing  $10^7$  counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases:  
15 Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate.  
20 Hybridization patterns are visualized using autoradiography or an alternative imaging means and compared.

#### **VII. Microarrays**

A chemical coupling procedure and an ink jet device can be used to synthesize array elements on the surface of a substrate. (See, e.g., Baldeschweiler, supra.) An array analogous to a  
25 dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced by hand or using available methods and machines and contain any appropriate number of elements. After hybridization, nonhybridized probes are removed and a scanner used to determine the levels and patterns of fluorescence. The degree of complementarity and the relative abundance of each probe  
30 which hybridizes to an element on the microarray may be assessed through analysis of the scanned images.

Full-length cDNAs, Expressed Sequence Tags (ESTs), or fragments thereof may comprise the elements of the microarray. Fragments suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). Full-length cDNAs, ESTs, or  
35 fragments thereof corresponding to one of the nucleotide sequences of the present invention, or

selected at random from a cDNA library relevant to the present invention, are arranged on an appropriate substrate, e.g., a glass slide. The cDNA is fixed to the slide using, e.g., UV cross-linking followed by thermal and chemical treatments and subsequent drying. (See, e.g., Schena, M. et al. (1995) *Science* 270:467-470; Shalon, D. et al. (1996) *Genome Res.* 6:639-645.) Fluorescent probes  
5 are prepared and used for hybridization to the elements on the substrate. The substrate is analyzed by procedures described above.

### VIII. Complementary Polynucleotides

Sequences complementary to the HRIP-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring HRIP. Although use of oligonucleotides  
10 comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of HRIP. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is  
15 designed to prevent ribosomal binding to the HRIP-encoding transcript.

### IX. Expression of HRIP

Expression and purification of HRIP is achieved using bacterial or virus-based expression systems. For expression of HRIP in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA  
20 transcription. Examples of such promoters include, but are not limited to, the *trp-lac (tac)* hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express HRIP upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of HRIP in eukaryotic cells is achieved by infecting insect  
25 or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding HRIP by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to  
30 infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) *Proc. Natl. Acad. Sci. USA* 91:3224-3227; Sandig, V. et al. (1996) *Hum. Gene Ther.* 7:1937-1945.)

In most expression systems, HRIP is synthesized as a fusion protein with, e.g., glutathione S-  
35 transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step,

affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from

5 HRIP at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, supra, ch. 10 and 16). Purified HRIP obtained by these methods can be used directly in the following

10 activity assay.

#### **X. Demonstration of HRIP Activity**

Kinase activity of HRIP is measured by the phosphorylation of appropriate substrates using gamma-labeled  $^{32}\text{P}$ -ATP and quantitation of the incorporated radioactivity using a beta radioisotope counter. HRIP is incubated with the protein substrate,  $^{32}\text{P}$ -ATP, and an appropriate kinase buffer.

15 The  $^{32}\text{P}$  incorporated into the product is separated from free  $^{32}\text{P}$ -ATP by electrophoresis and the incorporated  $^{32}\text{P}$  is counted. The amount of  $^{32}\text{P}$  recovered is proportional to the kinase activity of HRIP in the assay. A determination of the specific amino acid residue phosphorylated by protein kinase activity is made by phosphoamino acid analysis of the hydrolyzed protein.

Alternatively, protein phosphatase activity of HRIP is measured by the hydrolysis of P-nitrophenyl phosphate (PNPP). HRIP is incubated together with PNPP in HEPES buffer pH 7.5, in

20 the presence of 0.1% b-mercaptoethanol at 37°C for 60 min. The reaction is stopped by the addition of 6 ml of 10 N NaOH and the increase in light absorbance at 410 nm resulting from the hydrolysis of PNPP is measured using a spectrophotometer. The increase in light absorbance is proportional to the activity of HRIP in the assay (Diamond, R.H. et al. (1994) Mol. Cell Biol. 14:3752-3762).

#### **25 XI. Functional Assays**

HRIP function is assessed by expressing the sequences encoding HRIP at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include pCMV SPORT plasmid (Life Technologies) and pCR3.1 plasmid (Invitrogen), both of which

30 contain the cytomegalovirus promoter. 5-10  $\mu\text{g}$  of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2  $\mu\text{g}$  of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the

35 recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP;



Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of HRIP on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding HRIP and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding HRIP and other genes of interest can be analyzed by northern analysis or microarray techniques.

## **XII. Production of HRIP Specific Antibodies**

HRIP substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) *Methods Enzymol.* 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the HRIP amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, supra, ch. 11.)

Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (Perkin-Elmer) using fmoc-chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, supra.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide and anti-HRIP activity by, for example, binding the peptide or HRIP to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

**XIII. Purification of Naturally Occurring HRIP Using Specific Antibodies**

Naturally occurring or recombinant HRIP is substantially purified by immunoaffinity chromatography using antibodies specific for HRIP. An immunoaffinity column is constructed by covalently coupling anti-HRIP antibody to an activated chromatographic resin, such as  
5 CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing HRIP are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of HRIP (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt  
10 antibody/HRIP binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and HRIP is collected.

**XIV. Identification of Molecules Which Interact with HRIP**

HRIP, or biologically active fragments thereof, are labeled with <sup>125</sup>I Bolton-Hunter reagent. (See, e.g., Bolton A.E. and W.M. Hunter (1973) Biochem. J. 133:529-539.) Candidate molecules  
15 previously arrayed in the wells of a multi-well plate are incubated with the labeled HRIP, washed, and any wells with labeled HRIP complex are assayed. Data obtained using different concentrations of HRIP are used to calculate values for the number, affinity, and association of HRIP with the candidate molecules.

Alternatively, molecules interacting with HRIP are analyzed using the yeast two-hybrid  
20 system as described in Fields, S. and O. Song (1989, Nature 340:245-246), or using commercially available kits based on the two-hybrid system, such as the MATCHMAKER system (Clontech).

Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the  
25 invention. Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

Table 1

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
1	15	480457	LIVRBCT01	154793R6 (THP1PLB02), 480457H1 (LIVRBCT01), 480457X12R1 (LIVRBCT01), 480457X13R1 (LIVRBCT01), 1817968F6 (PROSNOT20), 3149829R6 (ADRENON04), 3766714H1 (BRSTNOT24)
2	16	563663	NEUTLPT01	285464X4 (EOSIHET02), 285464X8 (EOSIHET02), 563663H1 (NEUTLPT01)
3	17	1425842	BEPINON01	120376R6 (MUSCNOT01), 1425842H1 (BEPINON01), 1571293F1 (UTRSNOT05), 1851503F6 (LUNGFET03), 3596860F6 (FIBPNOT01)
4	18	2349047	COLSUCT01	1718442F6 (BLADNOT06), 1960909R6 (BRSTNOT04), 1960909T6 (BRSTNOT04), 2349047H1 (COLSUCT01), SBHA02478F1, SBHA02364F1, SBHA00749F1
5	19	2415617	HNT3AZT01	471426R6 (MMLR1DT01), 941244R1 (ADRENOT03), 1466986F1 (PANCTUT02), 1519153H1 (BLADTUT04), 2301362R6 (BRSTNOT05), 2415617H1 (HNT3AZT01), 2482778F6 (SMCANOT01), 4410796H1 (MONOTXT01)
6	20	3815186	TONSNOT03	1759852R3 (PITUNOT03), 3815186H1 (TONSNOT03), SBJA02796F1
7	21	5504544	BRABDIR01	046651X31 (CORNNOT01), 179527X7 (PLACNOB01), 905771T1 (COLNNOT08), 1291784F6 (PGANNOT03), 1291784T6 (PGANNOT03), 1542946X14 (PROSTUT04), 2759765R6 (THPIAZS08), 2908366F6 (THYMNOT05), 5504544H1 (BRABDIR01)
8	22	1511326	LUNGNOT14	1511326H1 (LUNGNOT14), 1511326F1 (LUNGNOT14), 1511326T1 (LUNGNOT14), 1511326T6 (LUNGNOT14), 2922438F6 (SININOT04), 3029015H1 (HEARFET02), 5167787H1 (MUSCDMT01)
9	23	1519120	BLADTUT04	1519120H1 (BLADTUT04), 1519120F6 (BLADTUT04), 1519120T6 (BLADTUT04)

Table 1 (cont.)

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
10	24	1673761	BLADNOT05	1673761H1 (BLADNOT05), 1673761F6 (BLADNOT05), 1673761T6 (BLADNOT05), 2024815R6 (KERANOT02), 2111282H1 (BRAITUT03), 2655266F6 (THYMNOT04), 4178671H1 (BRAINOT22), 5307169H1 (MONOTXT02), 1424569T1 (BEPINON01)
11	25	1270442	BRAINOT09	401269R6 (TMLR3DT01), 1270442H1 (BRAINOT09), 1672906F6 (BLADNOT05), 2596974H1 (OVARUT02), 2794801H1 (NPOLNOT01), 3289454H1 (BONRFET01), 3293681H1 (TLYJINT01), 3377437H1 (PENGNOT01), 3744502H1 (THYMNOT08)
12	26	1877133	LEUKNOT03	076555H1 (THP1PEB01), 1422142F6 (KIDNNOT09), 1877133H1 (LEUKNOT03), 2403383F6 (SMCANOT01), 2483377H1 (SMCANOT01)
13	27	2636759	BONTNOT01	2636759F6 (BONTNOT01), 2636759H1 (BONTNOT01), SBUA02427D1
14	28	2716815	THYRNOT09	078565R1 (SYNORAB01), 079174R1 (SYNORAB01), 1272190X14 (TESTTUT02), 1272190X23R1 (TESTTUT02), 1272190X24R1 (TESTTUT02), 2680827H1 (SINIUCT01), 2716815H1 (THYRNOT09), 2716815T6 (THYRNOT09), 3272956H1 (PROSEPT06)

Table 2

Protein SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences	Homologous Sequence	Analytical Methods
1	482	T160 S4 T211 S217 T313 T430 S52 S67 T140 S146 S201 S217 S224 S258 S275 S378 T467 T472	N57 N184 N353	Dual-specificity phosphatase catalytic site: E321-D461	Dual-specificity protein tyrosine phosphatase [Rattus norvegicus] (g1185552)	BLAST MOTIFS PFAM BLOCKS PRINTS
2	190	T35 T55 S131 S2 S183 Y147	N102	Protein kinase ATP-binding site: L39-K62 Protein kinase catalytic site: I154-L166 Eukaryotic protein kinase domain: I33-I186	DRAK2 kinase [Homo sapiens] (g3834356)	BLAST MOTIFS PFAM PRINTS
3	455	S252 S89 S234 S258 S268 S302 S342 S346 S364 S429 S434 S61 S96 S302 S410 T414 S415 Y343	N97 N159 N265 N409	Protein kinase ATP-binding site: V129-I141 Eukaryotic protein kinase domain: L16-I257 Leucine zipper: L294-L322	Serine/threonine protein kinase ZIP [Homo sapiens] (g561543)	BLAST MOTIFS PFAM PRINTS
4	485	S166 S283 S16 T167 S208 S242 T267 S283 T292 S306 T354 S278 T336 T370 S402 T412 S449 S483	N66 N400 N421 N481	Protein kinase catalytic site: L105 -L117 Eukaryotic protein kinase domain: R26-L247	Serine/threonine kinase RICK [Homo sapiens] (g3123887)	BLAST MOTIFS PRINTS PFAM
5	384	T130 T54 S181 T205 S371	N137	Diacylglycerol kinase catalytic domain: R16-L153	Sphingosine kinase [Mus musculus] (g3659694)	BLAST PFAM BLOCKS

Table 2 (cont.)

Protein SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences	Homologous Sequence	Analytical Methods
6	81	T37 S38 T52 T73 T54		Protein kinase C terminal domain: D33-F77	Protein kinase [Rattus norvegicus] (g206191)	BLAST PFAM
7	721	S40 S239 S640 T61 S68 S176 S196 S205 T241 S251 T416 T432 T655 T696 S49 T90 S230 T234 T235 S251 T255 T277 T416 S447 S484 S549 T696	N30 N274 N275 N297 N316 N572	Protein kinase ATP-binding site: L400-K423 Protein kinase catalytic site: I513-L525 Eukaryotic protein kinase domain: I394-L650 Phorbol ester/ diacylglycerol binding domain: H108-C157	Protein kinase C mu [Homo sapiens] (g438373)	BLAST MOTIFS PFAM PRINTS
8	249	S3 S4 T38 T137 S150 T64 T75 S107 S119 S196	N204	Tyrosine specific protein phosphatases active site: V88-I100 Tyrosine specific protein phosphatase: V88-S98	Putative tyrosine phosphatase [Homo sapiens] (g6650693)	MOTIFS BLOCKS BLAST
9	146	S125 S131		Eukaryotic protein kinase domain: Y12-L105	mCASK-A [Mus musculus] (g3087816)	MOTIFS BLAST PFAM
10	524	T21 T31 S77 S190 S237 S311 S511 S198 S207 T417 S440	N189	Eukaryotic protein kinase domain: P247-P492 Protein kinase signatures: L253-K276, L368-L380	Protein kinase homolog [Arabidopsis thaliana] (g2244835)	MOTIFS BLAST PFAM PRINTS PROFILES SCAN

Table 2 (cont.)

Protein SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences	Homologous Sequence	Analytical Methods
11	509	S224 T45 S80 T95 T216 T262 S307 S317 T325 S343 T388 S434 T452 S466 S7 S212 S282	N208	Protein phosphatase pp2a regulatory subunit, alternative splicing: S56-E462 Protein phosphatase pp2a regulatory subunit: S56-V312 EF-hand calcium-binding domain: D335-L347	Protein phosphatase 2A 72 kDa regulatory subunit [Homo sapiens] (g190222)	MOTIFS BLAST- GenBank BLAST- PRODOM BLAST-DOMO
12	142	S18 T26 T77 T4 T116	N114		TAK1 TGF-beta activated kinase [Xenopus laevis] (g3057036)	MOTIFS BLAST- GenBank
13	221	S21 S208 T2 T69 T170 S9 S16		Tyrosine specific protein phosphatase active site: V146-L158 Dual specificity protein tyrosine phosphatase: D121-L201 VH1-type dual specificity phosphatase: E57-V202	Tyrosine/serine phosphatase [Homo sapiens] (g181840)	MOTIFS BLAST- GenBank BLAST- PRODOM BLAST-DOMO HMMER-PFAM PROFILES SCAN
14	462	T119 S302 S390 S36 S59 S163 T167 T175 S177 T187 S218 S276 T309 S311 S434 T443 S35 S87 S200 T242 S334 S360 T436	N57	Tyrosine kinase: L139-P427	Multihormonally regulated gene [Rattus norvegicus] (g1561667)	MOTIFS BLAST- GenBank BLAST-DOMO

Table 3

Polynucleotide SEQ ID NO:	Useful Fragments	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
15	435-479	Reproductive (0.300) Cardiovascular (0.150) Hematopoietic/Immune (0.150)	Cell Proliferation (0.450) Inflammation (0.500)	PBLUESCRIPT
16	219-263	Hematopoietic/Immune (0.455) Cardiovascular (0.182) Developmental (0.091)	Cell Proliferation (0.546) Inflammation (0.455)	PBLUESCRIPT
17	1191-1235	Reproductive (0.250) Cardiovascular (0.214) Gastrointestinal (0.214)	Cell Proliferation (0.465) Inflammation (0.322)	pT7T3
18	542-586 974-1018	Gastrointestinal (0.391) Hematopoietic/Immune (0.174) Reproductive (0.174)	Inflammation (0.478) Cell Proliferation (0.391)	pINCY
19	217-261 718-761	Cardiovascular (0.256) Reproductive (0.179) Hematopoietic/Immune (0.154)	Cell Proliferation (0.564) Inflammation (0.334)	pINCY
20	380-424	Nervous (0.520) Hematopoietic/Immune (0.240) Reproductive (0.120)	Inflammation (0.520) Cell Proliferation (0.400)	pINCY
21	487-531 1027-1071	Reproductive (0.296) Hematopoietic/Immune (0.198) Gastrointestinal (0.111)	Cell Proliferation (0.568) Inflammation (0.358)	pINCY
22	379-423	Cardiovascular (0.267) Gastrointestinal (0.200) Hematopoietic/Immune (0.200)	Inflammation (0.333) Cancer (0.267) Trauma (0.200)	pINCY
23	192-218	Nervous (0.500) Urologic (0.500)	Cancer (0.500) Cell Line (0.500)	pINCY



Table 3 (cont.)

Polynucleotide SEQ ID NO:	Useful Fragments	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
24	163-207	Nervous (0.333) Cardiovascular (0.133) Dermatologic (0.133)	Cancer (0.400) Fetal/Cell Line (0.200) Inflammation (0.133)	pINCY
25	19-63	Cardiovascular (0.182) Hematopoietic/Immune (0.182) Reproductive (0.182)	Cancer (0.515) Cell Proliferation (0.242) Inflammation (0.242)	pINCY
26	297-343	Cardiovascular (0.250) Hematopoietic/Immune (0.150) Musculoskeletal (0.150)	Cancer (0.300) Cell Proliferation (0.250) Inflammation (0.200)	pINCY
27	271-315	Endocrine (0.500) Musculoskeletal (0.500)	Cancer (0.500)	pINCY
28	161-207	Reproductive (0.241) Gastrointestinal (0.233) Cardiovascular (0.150)	Cancer (0.429) Inflammation (0.263) Cell Proliferation (0.211)	pINCY

Table 4

Polynucleotide SEQ ID NO:	Library	Library Comment
15	LIVRBCT01	Library was constructed using RNA isolated from the liver tissue of a patient with primary biliary cirrhosis who had a liver transplant.
16	NEUTLPT01	Library was constructed using RNA isolated from peripheral blood granulocytes collected by density gradient centrifugation through Ficoll-Hypaque. The cells were isolated from buffy coat units obtained from unrelated male and female donors. Cells were cultured in 100 ng/ml E. coli LPS for 30 minutes.
17	BEPINON01	This normalized bronchial epithelium library was constructed from 5.12 million independent clones from a bronchial epithelium library. RNA was made from a bronchial epithelium primary cell line derived from a 54-year-old Caucasian male. The normalization and hybridization conditions were adapted from Soares et al., Proc. Natl. Acad. Sci. USA (1994) 91:9228, using a longer (24-hour) reannealing hybridization period.
18	COLSUCT01	Library was constructed using RNA isolated from diseased sigmoid colon tissue obtained from a 70-year-old Caucasian male during colectomy with permanent ileostomy. Pathology indicated chronic ulcerative colitis. Patient history included benign neoplasm of the colon. Family history included atherosclerotic coronary artery disease and myocardial infarctions.
19	HNT3AZT01	Library was constructed using RNA isolated from the hNT2 cell line (derived from a human teratocarcinoma that exhibited properties characteristic of a committed neuronal precursor). Cells were treated for three days with 0.35 micromolar 5-aza-2'-deoxycytidine (AZ).
20	TONSNOT03	Library was constructed using RNA isolated from diseased left tonsil tissue removed from a 6-year-old Caucasian male during adenotonsillectomy. Pathology indicated reactive lymphoid hyperplasia, bilaterally. Family history included benign hypertension, myocardial infarction, and atherosclerotic coronary artery disease.
21	BRABDIR01	Library was constructed using RNA isolated from diseased cerebellum tissue removed from the brain of a 57-year-old Caucasian male, who died from a cerebrovascular accident.
22	LUNGNOT14	Library was constructed using RNA isolated from lung tissue removed from the left lower lobe of a 47-year-old Caucasian male during a segmental lung resection. Pathology for the associated tumor tissue indicated a grade 4 adenocarcinoma, and the parenchyma showed calcified granuloma. Patient history included benign hypertension and chronic obstructive pulmonary disease. Family history included type II diabetes and acute myocardial infarction.

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Library	Library Comment
23	BLADTUT04	Library was constructed using RNA isolated from bladder tumor tissue removed from a 60-year-old Caucasian male during a radical cystectomy, prostatectomy, and vasectomy. Pathology indicated grade 3 transitional cell carcinoma in the left bladder wall. Carcinoma in-situ was identified in the dome and trigone. Family history included type I diabetes, malignant neoplasm of the stomach, atherosclerotic coronary artery disease, and acute myocardial infarction.
24	BLADNOT05	Library was constructed using RNA isolated from bladder tissue removed from a 60-year-old Caucasian male during a radical cystectomy, prostatectomy, and vasectomy. Pathology for the associated tumor tissue indicated grade 3 transitional cell carcinoma. Carcinoma in-situ was identified in the dome and trigone. Family history included type I diabetes, malignant neoplasm of the stomach, atherosclerotic coronary artery disease, and acute myocardial infarction.
25	BRAINOT09	Library was constructed using RNA isolated from brain tissue removed from a Caucasian male fetus, who died at 23 weeks' gestation.
26	LEUKNOT03	Library was constructed using RNA isolated from white blood cells of a 27-year-old female with blood type A+. The donor tested negative for cytomegalovirus (CMV).
27	BONTNOT01	Library was constructed using RNA isolated from tibial periosteum removed from a 20-year-old Caucasian male during a hemipelvectomy with amputation above the knee. Pathology for the associated tumor tissue indicated partially necrotic and cystic osteoblastic grade 3 osteosarcoma (post-chemotherapy). Family history included osteogenesis imperfecta, closed fracture, and type II diabetes.
28	THYRNOT09	Library was constructed using RNA isolated from diseased thyroid tissue removed from an 18-year-old Caucasian female during an unilateral thyroid lobectomy and regional lymph node excision. Pathology indicated adenomatous goiter. This was associated with a follicular adenoma of the thyroid. Family history included thyroid cancer.

Table 5

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25: 3389-3402.	ESTs: Probability value= 1.0E-8 or less Full Length sequences: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises at least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad. Sci. 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183: 63-98; and Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489.	ESTs: fasta E value=1.06E-6 Assembled ESTs: fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value= 1.0E-8 or less Full Length sequences: fastx score= 100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S and J.G. Henikoff, Nucl. Acid Res., 19:6565-72, 1991. J.G. Henikoff and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37: 417-424.	Score=1000 or greater; Ratio of Score/Strength = 0.75 or larger; and, if applicable, Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) J. Mol. Biol., 235:1501-1531; Sonnhammer, E.L.L. et al. (1998) Nucleic Acids Res. 26:320-322.	Score=10-50 bits for PFAM hits, depending on individual protein families

Table 5 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25: 217-221.	Normalized quality score $\geq$ GCG-specified "HIGH" value for that particular Prosite motif. Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M. S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12: 431-439.	Score=3.5 or greater
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch et al. <i>supra</i> ; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

What is claimed is:

1. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of:
  - 5 a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-14,
  - b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-14,
  - c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-14, and
  - 10 d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-14.
2. An isolated polypeptide of claim 1 selected from the group consisting of SEQ ID NO:1-14.
- 15 3. An isolated polynucleotide encoding a polypeptide of claim 1.
4. An isolated polynucleotide of claim 3 selected from the group consisting of SEQ ID NO:15-28.
- 20 5. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.
6. A cell transformed with a recombinant polynucleotide of claim 5.
- 25 7. A transgenic organism comprising a recombinant polynucleotide of claim 5.
8. A method for producing a polypeptide of claim 1, the method comprising:
  - a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said
  - 30 cell is transformed with a recombinant polynucleotide, and said recombinant polynucleotide comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and
  - b) recovering the polypeptide so expressed.
- 35 9. An isolated antibody which specifically binds to a polypeptide of claim 1.

10. An isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of:
- a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:15-28,
  - b) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:15-28,
  - c) a polynucleotide sequence complementary to a),
  - d) a polynucleotide sequence complementary to b), and
  - e) an RNA equivalent of a)-d).
11. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 10.
12. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 10, the method comprising:
- a) hybridizing the sample with a probe comprising at least 16 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide, and
  - b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.
13. A method of claim 12, wherein the probe comprises at least 30 contiguous nucleotides.
14. A method of claim 12, wherein the probe comprises at least 60 contiguous nucleotides.
15. A pharmaceutical composition comprising an effective amount of a polypeptide of claim 1 and a pharmaceutically acceptable excipient.
16. A method for treating a disease or condition associated with decreased expression of functional HRIP, comprising administering to a patient in need of such treatment the pharmaceutical composition of claim 15.
17. A method for screening a compound for effectiveness as an agonist of a polypeptide of claim 1, the method comprising:
- a) exposing a sample comprising a polypeptide of claim 1 to a compound, and

b) detecting agonist activity in the sample.

18. A pharmaceutical composition comprising an agonist compound identified by a method of claim 17 and a pharmaceutically acceptable excipient.

5

19. A method for treating a disease or condition associated with decreased expression of functional HRIP, comprising administering to a patient in need of such treatment a pharmaceutical composition of claim 18.

10 20. A method for screening a compound for effectiveness as an antagonist of a polypeptide of claim 1, the method comprising:

- a) exposing a sample comprising a polypeptide of claim 1 to a compound, and
- b) detecting antagonist activity in the sample.

15 21. A pharmaceutical composition comprising an antagonist compound identified by a method of claim 20 and a pharmaceutically acceptable excipient.

20 22. A method for treating a disease or condition associated with overexpression of functional HRIP, comprising administering to a patient in need of such treatment a pharmaceutical composition of claim 21.

23. A method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence of claim 4, the method comprising:

- 25
- a) exposing a sample comprising the target polynucleotide to a compound, and
  - b) detecting altered expression of the target polynucleotide.



PF-0683 PCT

## SEQUENCE LISTING

&lt;110&gt; INCYTE PHARMACEUTICALS, INC.

BANDMAN, Olga

TANG, Y. Tom

YUE, Henry

HILLMAN, Jennifer L.

BAUGHN, Mariah R.

AZIMZAI, Yalda

LU, Dyung Aina M.

AU-YOUNG, Janice

&lt;120&gt; REGULATORS OF INTRACELLULAR PHOSPHORYLATION

&lt;130&gt; PF-0683 PCT

&lt;140&gt; To Be Assigned

&lt;141&gt; Herewith

&lt;150&gt; 60/125,593; 60/135,049; 60/143,188

&lt;151&gt; 1999-03-18; 1999-05-20; 1999-07-09

&lt;160&gt; 28

&lt;170&gt; PERL Program

&lt;210&gt; 1

&lt;211&gt; 482

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 480457CD1

&lt;400&gt; 1

Met	Pro	Pro	Ser	Pro	Leu	Asp	Asp	Arg	Val	Val	Val	Ala	Leu	Ser	1	5	10	15
Arg	Pro	Val	Arg	Pro	Gln	Asp	Leu	Asn	Leu	Cys	Leu	Asp	Ser	Ser	20	25	30	35
Tyr	Leu	Gly	Ser	Ala	Asn	Pro	Gly	Ser	Asn	Ser	His	Pro	Pro	Val	40	45	50	55
Ile	Ala	Thr	Thr	Val	Val	Ser	Leu	Lys	Ala	Ala	Asn	Leu	Thr	Tyr	60	65	70	75
Met	Pro	Ser	Ser	Ser	Gly	Ser	Ala	Arg	Ser	Leu	Asn	Cys	Gly	Cys	80	85	90	95
Ser	Ser	Ala	Ser	Cys	Cys	Thr	Val	Ala	Thr	Tyr	Asp	Lys	Asp	Asn	100	105	110	115
Gln	Ala	Gln	Thr	Gln	Ala	Ile	Ala	Ala	Gly	Thr	Thr	Thr	Thr	Ala	120	125	130	135
Ile	Gly	Thr	Ser	Thr	Thr	Cys	Pro	Ala	Asn	Gln	Met	Val	Asn	Asn	140	145	150	155
Asn	Glu	Asn	Thr	Gly	Ser	Leu	Ser	Pro	Ser	Ser	Gly	Val	Gly	Ser	160	165	170	175
Pro	Val	Ser	Gly	Thr	Pro	Lys	Gln	Leu	Ala	Ser	Ile	Lys	Ile	Ile				
Tyr	Pro	Asn	Asp	Leu	Ala	Lys	Lys	Met	Thr	Lys	Cys	Ser	Lys	Ser				
His	Leu	Pro	Ser	Gln	Gly	Pro	Val	Ile	Ile	Asp	Cys	Arg	Pro	Phe				

PF-0683 PCT

Met	Glu	Tyr	Asn	Lys	Ser	His	Ile	Gln	Gly	Ala	Val	His	Ile	Asn	
				185					190					195	
Cys	Ala	Asp	Lys	Ile	Ser	Arg	Arg	Arg	Leu	Gln	Gln	Gly	Lys	Ile	
				200					205					210	
Thr	Val	Leu	Asp	Leu	Ile	Ser	Cys	Arg	Glu	Gly	Lys	Asp	Ser	Phe	
				215					220					225	
Lys	Arg	Ile	Phe	Ser	Lys	Glu	Ile	Ile	Val	Tyr	Asp	Glu	Asn	Thr	
				230					235					240	
Asn	Glu	Pro	Ser	Arg	Val	Met	Pro	Ser	Gln	Pro	Leu	His	Ile	Val	
				245					250					255	
Leu	Glu	Ser	Leu	Lys	Arg	Glu	Gly	Lys	Glu	Pro	Leu	Val	Leu	Lys	
				260					265					270	
Gly	Gly	Leu	Ser	Ser	Phe	Lys	Gln	Asn	His	Glu	Asn	Leu	Cys	Asp	
				275					280					285	
Asn	Ser	Leu	Gln	Leu	Gln	Glu	Cys	Arg	Glu	Val	Gly	Gly	Gly	Ala	
				290					295					300	
Ser	Ala	Ala	Ser	Ser	Leu	Leu	Pro	Gln	Pro	Ile	Pro	Thr	Thr	Pro	
				305					310					315	
Asp	Ile	Glu	Asn	Ala	Glu	Leu	Thr	Pro	Ile	Leu	Pro	Phe	Leu	Phe	
				320					325					330	
Leu	Gly	Asn	Glu	Gln	Asp	Ala	Gln	Asp	Leu	Asp	Thr	Met	Gln	Arg	
				335					340					345	
Leu	Asn	Ile	Gly	Tyr	Val	Ile	Asn	Val	Thr	Thr	His	Leu	Pro	Leu	
				350					355					360	
Tyr	His	Tyr	Glu	Lys	Gly	Leu	Phe	Asn	Tyr	Lys	Arg	Leu	Pro	Ala	
				365					370					375	
Thr	Asp	Ser	Asn	Lys	Gln	Asn	Leu	Arg	Gln	Tyr	Phe	Glu	Glu	Ala	
				380					385					390	
Phe	Glu	Phe	Ile	Glu	Glu	Ala	His	Gln	Cys	Gly	Lys	Gly	Leu	Leu	
				395					400					405	
Ile	His	Cys	Gln	Ala	Gly	Val	Ser	Arg	Ser	Ala	Thr	Ile	Val	Ile	
				410					415					420	
Ala	Tyr	Leu	Met	Lys	His	Thr	Arg	Met	Thr	Met	Thr	Asp	Ala	Tyr	
				425					430					435	
Lys	Phe	Val	Lys	Gly	Lys	Arg	Pro	Ile	Ile	Ser	Pro	Asn	Leu	Asn	
				440					445					450	
Phe	Met	Gly	Gln	Leu	Leu	Glu	Phe	Glu	Glu	Asp	Leu	Asn	Asn	Gly	
				455					460					465	
Val	Thr	Pro	Arg	Ile	Leu	Thr	Pro	Lys	Leu	Met	Gly	Val	Glu	Thr	
				470					475					480	
Val	Val														

&lt;210&gt; 2

&lt;211&gt; 190

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 563663CD1

&lt;400&gt; 2

Met	Ser	Arg	Arg	Arg	Phe	Asp	Cys	Arg	Ser	Ile	Ser	Gly	Leu	Leu	
1				5					10					15	
Thr	Thr	Thr	Pro	Gln	Ile	Pro	Ile	Lys	Met	Glu	Asn	Phe	Asn	Asn	
				20					25					30	
Phe	Tyr	Ile	Leu	Thr	Ser	Lys	Glu	Leu	Gly	Arg	Gly	Lys	Phe	Ala	
				35					40					45	
Val	Val	Arg	Gln	Cys	Ile	Ser	Lys	Ser	Thr	Gly	Gln	Glu	Tyr	Ala	
				50					55					60	

PF-0683 PCT

Ala	Lys	Phe	Leu	Lys	Lys	Arg	Arg	Arg	Gly	Gln	Asp	Cys	Arg	Ala	
				65					70					75	
Glu	Ile	Leu	His	Glu	Ile	Ala	Val	Leu	Glu	Leu	Ala	Lys	Ser	Cys	
				80					85					90	
Pro	Arg	Val	Ile	Asn	Leu	His	Glu	Val	Tyr	Glu	Asn	Thr	Ser	Glu	
				95					100					105	
Ile	Ile	Leu	Ile	Leu	Glu	Tyr	Ala	Ala	Gly	Glu	Ile	Phe	Ser		
				110					115					120	
Leu	Cys	Leu	Pro	Glu	Leu	Ala	Glu	Met	Val	Ser	Glu	Asn	Asp	Val	
				125					130					135	
Ile	Arg	Leu	Ile	Lys	Gln	Ile	Leu	Glu	Gly	Val	Tyr	Tyr	Leu	His	
				140					145					150	
Gln	Asn	Asn	Ile	Val	His	Leu	Asp	Leu	Lys	Pro	Gln	Asn	Ile	Leu	
				155					160					165	
Leu	Ser	Ser	Ile	Tyr	Pro	Leu	Gly	Asp	Ile	Lys	Ile	Val	Asp	Gly	
				170					175					180	
Gly	Met	Ser	Arg	Lys	Ile	Gly	Gln	Cys	Val						
				185					190						

&lt;210&gt; 3

&lt;211&gt; 455

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 1425842CD1

&lt;400&gt; 3

Met	Ser	Ser	Leu	Gly	Ala	Ser	Phe	Val	Gln	Ile	Lys	Phe	Asp	Asp	
1				5					10					15	
Leu	Gln	Phe	Phe	Glu	Asn	Cys	Gly	Gly	Gly	Ser	Phe	Gly	Ser	Val	
				20					25					30	
Tyr	Arg	Ala	Lys	Trp	Ile	Ser	Gln	Asp	Lys	Glu	Val	Ala	Val	Lys	
				35					40					45	
Lys	Leu	Leu	Lys	Ile	Glu	Lys	Glu	Ala	Glu	Ile	Leu	Ser	Val	Leu	
				50					55					60	
Ser	His	Arg	Asn	Ile	Ile	Gln	Phe	Tyr	Gly	Val	Ile	Leu	Glu	Pro	
				65					70					75	
Pro	Asn	Tyr	Gly	Ile	Val	Thr	Glu	Tyr	Ala	Ser	Leu	Gly	Ser	Leu	
				80					85					90	
Tyr	Asp	Tyr	Ile	Asn	Ser	Asn	Arg	Ser	Glu	Glu	Met	Asp	Met	Asp	
				95					100					105	
His	Ile	Met	Thr	Trp	Ala	Thr	Asp	Val	Ala	Lys	Gly	Met	His	Tyr	
				110					115					120	
Leu	His	Met	Glu	Ala	Pro	Val	Lys	Val	Ile	His	Arg	Asp	Leu	Lys	
				125					130					135	
Ser	Arg	Asn	Val	Val	Ile	Ala	Ala	Asp	Gly	Val	Leu	Lys	Ile	Cys	
				140					145					150	
Asp	Phe	Gly	Ala	Ser	Arg	Phe	His	Asn	His	Thr	Thr	His	Met	Ser	
				155					160					165	
Leu	Val	Gly	Thr	Phe	Pro	Trp	Met	Ala	Pro	Glu	Val	Ile	Gln	Ser	
				170					175					180	
Leu	Pro	Val	Ser	Glu	Thr	Cys	Asp	Thr	Tyr	Ser	Tyr	Gly	Val	Val	
				185					190					195	
Leu	Trp	Glu	Met	Leu	Thr	Arg	Glu	Val	Pro	Phe	Lys	Gly	Leu	Glu	
				200					205					210	
Gly	Leu	Gln	Val	Ala	Trp	Leu	Val	Val	Glu	Lys	Asn	Glu	Arg	Leu	
				215					220					225	
Thr	Ile	Pro	Ser	Ser	Cys	Pro	Arg	Ser	Phe	Ala	Glu	Leu	Leu	His	

PF-0683 PCT

	230		235		240
Gln Cys Trp Glu	Ala Asp Ala Lys Lys	Arg Pro Ser Phe Lys	Gln		
	245		250		255
Ile Ile Ser Ile	Leu Glu Ser Met Ser	Asn Asp Thr Ser Leu	Pro		
	260		265		270
Asp Lys Cys Asn	Ser Phe Leu His Asn	Lys Ala Glu Trp Arg	Cys		
	275		280		285
Glu Ile Glu Ala	Thr Leu Glu Arg Leu	Lys Lys Leu Glu Arg	Asp		
	290		295		300
Leu Ser Phe Lys	Glu Gln Glu Leu Lys	Glu Arg Glu Arg Arg	Leu		
	305		310		315
Lys Met Trp Glu	Gln Lys Leu Thr Glu	Gln Ser Asn Thr Pro	Leu		
	320		325		330
Leu Leu Pro Leu	Ala Ala Arg Met Ser	Glu Glu Ser Tyr Phe	Glu		
	335		340		345
Ser Lys Thr Glu	Glu Ser Asn Ser Ala	Glu Met Ser Cys Gln	Ile		
	350		355		360
Thr Ala Thr Ser	Asn Gly Glu Gly His	Gly Met Asn Pro Ser	Leu		
	365		370		375
Gln Ala Met Met	Leu Met Gly Phe Gly	Asp Ile Phe Ser Met	Asn		
	380		385		390
Lys Ala Gly Ala	Val Met His Ser Gly	Met Gln Ile Asn Met	Gln		
	395		400		405
Ala Lys Gln Asn	Ser Ser Lys Thr Thr	Ser Lys Arg Arg Gly	Lys		
	410		415		420
Lys Val Asn Met	Ala Leu Gly Phe Ser	Asp Phe Asp Leu Ser	Glu		
	425		430		435
Gly Asp Asp Asp	Asp Asp Asp Asp Gly	Glu Glu Glu Asp Asn	Asp		
	440		445		450
Met <del>Asp</del> Asn Ser	Glu				455
	455				

&lt;210&gt; 4

&lt;211&gt; 485

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 2349047CD1

&lt;400&gt; 4

Met Ala Lys Gly Gly	Ile Phe Pro Arg	Pro Arg Cys Asp Ser	Ser
1	5	10	15
Ser Leu Leu Glu Cys	Arg Lys Ala Ile	Ser Arg Glu Val Lys	Ala
	20	25	30
Met Ala Ser Leu Asp	Asn Glu Phe Val	Leu Arg Leu Glu Gly	Val
	35	40	45
Ile Glu Lys Val Asn	Trp Asp Gln Asp	Pro Lys Pro Ala Leu	Val
	50	55	60
Thr Lys Phe Met Glu	Asn Gly Ser Leu	Ser Gly Leu Leu Gln	Ser
	65	70	75
Gln Cys Pro Arg Pro	Trp Pro Leu Leu	Cys Arg Leu Leu Lys	Glu
	80	85	90
Val Val Leu Gly Met	Phe Tyr Leu His	Asp Gln Asn Pro Val	Leu
	95	100	105
Leu His Arg Asp Leu	Lys Pro Ser Asn	Val Leu Leu Asp Pro	Glu
	110	115	120
Leu His Val Lys Leu	Ala Asp Phe Gly	Leu Ser Thr Phe Gln	Gly
	125	130	135

PF-0683 PCT

Gly	Ser	Gln	Ser	Gly	Thr	Gly	Ser	Gly	Glu	Pro	Gly	Gly	Thr	Leu
				140					145					150
Gly	Tyr	Leu	Ala	Pro	Glu	Leu	Phe	Val	Asn	Val	Asn	Arg	Lys	Ala
				155					160					165
Ser	Thr	Ala	Ser	Asp	Val	Tyr	Ser	Phe	Gly	Ile	Leu	Met	Trp	Ala
				170					175					180
Val	Leu	Ala	Gly	Arg	Glu	Val	Glu	Leu	Pro	Thr	Glu	Pro	Ser	Leu
				185					190					195
Val	Tyr	Glu	Ala	Val	Cys	Asn	Arg	Gln	Asn	Arg	Pro	Ser	Leu	Ala
				200					205					210
Glu	Leu	Pro	Gln	Ala	Gly	Pro	Glu	Thr	Pro	Gly	Leu	Glu	Gly	Leu
				215					220					225
Lys	Glu	Leu	Met	Gln	Leu	Cys	Trp	Ser	Ser	Glu	Pro	Lys	Asp	Arg
				230					235					240
Pro	Ser	Phe	Gln	Glu	Cys	Leu	Pro	Lys	Thr	Asp	Glu	Val	Phe	Gln
				245					250					255
Met	Val	Glu	Asn	Asn	Met	Asn	Ala	Ala	Val	Ser	Thr	Val	Lys	Asp
				260					265					270
Phe	Leu	Ser	Gln	Leu	Arg	Ser	Ser	Asn	Arg	Arg	Phe	Ser	Ile	Pro
				275					280					285
Glu	Ser	Gly	Gln	Gly	Gly	Thr	Glu	Met	Asp	Gly	Phe	Arg	Arg	Thr
				290					295					300
Ile	Glu	Asn	Gln	His	Ser	Arg	Asn	Asp	Val	Met	Val	Ser	Glu	Trp
				305					310					315
Leu	Asn	Lys	Leu	Asn	Leu	Glu	Glu	Pro	Pro	Ser	Ser	Val	Pro	Lys
				320					325					330
Lys	Cys	Pro	Ser	Leu	Thr	Lys	Arg	Ser	Arg	Ala	Gln	Glu	Glu	Gln
				335					340					345
Val	Pro	Gln	Ala	Trp	Thr	Ala	Gly	Thr	Ser	Ser	Asp	Ser	Met	Ala
				350					355					360
Gln	Pro	Pro	Gln	Thr	Pro	Glu	Thr	Ser	Thr	Phe	Arg	Asn	Gln	Met
				365					370					375
Pro	Ser	Pro	Thr	Ser	Thr	Gly	Thr	Pro	Ser	Pro	Gly	Pro	Arg	Gly
				380					385					390
Asn	Gln	Gly	Ala	Glu	Arg	Gln	Gly	Met	Asn	Trp	Ser	Cys	Arg	Thr
				395					400					405
Pro	Glu	Pro	Asn	Pro	Val	Thr	Gly	Arg	Pro	Leu	Val	Asn	Ile	Tyr
				410					415					420
Asn	Cys	Ser	Gly	Val	Gln	Val	Gly	Asp	Asn	Asn	Tyr	Leu	Thr	Met
				425					430					435
Gln	Gln	Thr	Thr	Ala	Leu	Pro	Thr	Trp	Gly	Leu	Ala	Pro	Ser	Gly
				440					445					450
Lys	Gly	Arg	Gly	Leu	Gln	His	Pro	Pro	Pro	Val	Gly	Ser	Gln	Glu
				455					460					465
Gly	Pro	Lys	Asp	Pro	Glu	Ala	Trp	Ser	Arg	Pro	Gln	Gly	Trp	Tyr
				470					475					480
Asn	His	Ser	Gly	Lys										
				485										

&lt;210&gt; 5

&lt;211&gt; 384

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 2415617CD1

&lt;400&gt; 5

Met Asp Pro Ala Gly Gly Pro Arg Gly Val Leu Pro Arg Pro Cys

PF-0683 PCT

1	5	10	15
Arg Val Leu Val	Leu Leu Asn Pro Arg	Gly Gly Lys Gly Lys	Ala
20	25	30	
Leu Gln Leu Phe	Arg Ser His Val Gln	Pro Leu Leu Ala Glu	Ala
35	40	45	
Glu Ile Ser Phe	Thr Leu Met Leu Thr	Glu Arg Arg Asn His	Ala
50	55	60	
Arg Glu Leu Val	Arg Ser Glu Glu Leu	Gly Arg Trp Asp Ala	Leu
65	70	75	
Val Val Met Ser	Gly Asp Gly Leu Met	His Glu Val Val Asn	Gly
80	85	90	
Leu Met Glu Arg	Pro Asp Trp Glu Thr	Ala Ile Gln Lys Pro	Leu
95	100	105	
Cys Ser Leu Pro	Ala Gly Ser Gly Asn	Ala Leu Ala Ala Ser	Leu
110	115	120	
Asn His Tyr Ala	Gly Tyr Glu Gln Val	Thr Asn Glu Asp Leu	Leu
125	130	135	
Thr Asn Cys Thr	Leu Leu Leu Cys Arg	Arg Leu Leu Ser Pro	Met
140	145	150	
Asn Leu Leu Ser	Leu His Thr Ala Ser	Gly Leu Arg Leu Phe	Ser
155	160	165	
Val Leu Ser Leu	Ala Trp Gly Phe Ile	Ala Asp Val Asp Leu	Glu
170	175	180	
Ser Glu Lys Tyr	Arg Arg Leu Gly Glu	Met Arg Phe Thr Leu	Gly
185	190	195	
Thr Phe Leu Arg	Leu Ala Ala Leu Arg	Thr Tyr Arg Gly Arg	Leu
200	205	210	
Ala Tyr Leu Pro	Val Gly Arg Val Gly	Ser Lys Thr Pro Ala	Ser
215	220	225	
Pro Val Val Val	Gln Gln Gly Pro Val	Asp Ala His Leu Val	Pro
230	235	240	
Leu Glu Glu Pro	Val Pro Ser His Trp	Thr Val Val Pro Asp	Glu
245	250	255	
Asp Phe Val Leu	Val Leu Ala Leu Leu	His Ser His Leu Gly	Ser
260	265	270	
Glu Met Phe Ala	Ala Pro Met Gly Arg	Cys Ala Ala Gly Val	Met
275	280	285	
His Leu Phe Tyr	Val Arg Ala Gly Val	Ser Arg Ala Met Leu	Leu
290	295	300	
Arg Leu Phe Leu	Ala Met Glu Lys Gly	Arg His Met Glu Tyr	Glu
305	310	315	
Cys Pro Tyr Leu	Val Tyr Val Pro Val	Val Ala Phe Arg Leu	Glu
320	325	330	
Pro Lys Asp Gly	Lys Gly Val Phe Ala	Val Asp Gly Glu Leu	Met
335	340	345	
Val Ser Glu Ala	Val Gln Gly Gln Val	His Pro Asn Tyr Phe	Trp
350	355	360	
Met Val Ser Gly	Cys Val Glu Pro Pro	Pro Ser Trp Lys Pro	Gln
365	370	375	
Gln Met Pro Pro	Pro Glu Glu Pro Leu		
380			

<210> 6  
 <211> 81  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 3815186CD1

PF-0683 PCT

&lt;400&gt; 6

Met	Arg	Trp	Tyr	Gln	Pro	Pro	Asn	Asp	Trp	Arg	Ile	Leu	Val	Leu	
1				5					10					15	
Cys	Leu	Ser	Ser	Tyr	Ala	Val	Leu	Met	Cys	Leu	Leu	Ser	Ile	Trp	
				20					25					30	
Gln	Arg	Asp	Lys	Arg	Asp	Thr	Ser	Asn	Phe	Asp	Lys	Glu	Phe	Thr	
				35					40					45	
Arg	Gln	Pro	Val	Glu	Leu	Thr	Pro	Thr	Asp	Lys	Leu	Phe	Ile	Met	
				50					55					60	
Asn	Leu	Asp	Gln	Asn	Glu	Phe	Ala	Gly	Phe	Ser	Tyr	Thr	Asn	Pro	
				65					70					75	
Glu	Phe	Val	Ile	Asn	Val										
				80											

&lt;210&gt; 7

&lt;211&gt; 721

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 5504544CD1

&lt;400&gt; 7

Met	Leu	Phe	Gly	Leu	Val	Arg	Gln	Gly	Leu	Lys	Cys	Asp	Gly	Cys	
1				5					10					15	
Gly	Leu	Asn	Tyr	His	Lys	Arg	Cys	Ala	Phe	Ser	Ile	Pro	Asn	Asn	
				20					25					30	
Cys	Ser	Gly	Ala	Arg	Lys	Arg	Arg	Leu	Ser	Ser	Thr	Ser	Leu	Ala	
				35					40					45	
Ser	Gly	His	Ser	Val	Arg	Leu	Gly	Thr	Ser	Glu	Ser	Leu	Pro	Cys	
				50					55					60	
Thr	Ala	Glu	Glu	Leu	Ser	Arg	Ser	Thr	Thr	Glu	Leu	Leu	Pro	Arg	
				65					70					75	
Arg	Pro	Pro	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ala	Ser	Ser	Tyr	Thr	
				80					85					90	
Gly	Arg	Pro	Ile	Glu	Leu	Asp	Lys	Met	Leu	Leu	Ser	Lys	Val	Lys	
				95					100					105	
Val	Pro	His	Thr	Phe	Leu	Ile	His	Ser	Tyr	Thr	Arg	Pro	Thr	Val	
				110					115					120	
Cys	Gln	Ala	Cys	Lys	Lys	Leu	Leu	Lys	Gly	Leu	Phe	Arg	Gln	Gly	
				125					130					135	
Leu	Gln	Cys	Lys	Asp	Cys	Lys	Phe	Asn	Cys	His	Lys	Arg	Cys	Ala	
				140					145					150	
Thr	Arg	Val	Pro	Asn	Asp	Cys	Leu	Gly	Glu	Ala	Leu	Ile	Asn	Gly	
				155					160					165	
Asp	Val	Pro	Met	Glu	Glu	Ala	Thr	Asp	Phe	Ser	Glu	Ala	Asp	Lys	
				170					175					180	
Ser	Ala	Leu	Met	Asp	Glu	Ser	Glu	Asp	Ser	Gly	Val	Ile	Pro	Gly	
				185					190					195	
Ser	His	Ser	Glu	Asn	Ala	Leu	His	Ala	Ser	Glu	Glu	Glu	Glu	Gly	
				200					205					210	
Glu	Gly	Gly	Lys	Ala	Gln	Ser	Ser	Leu	Gly	Tyr	Ile	Pro	Leu	Met	
				215					220					225	
Arg	Val	Val	Gln	Ser	Val	Arg	His	Thr	Thr	Arg	Lys	Ser	Ser	Thr	
				230					235					240	
Thr	Leu	Arg	Glu	Gly	Trp	Val	Val	His	Tyr	Ser	Asn	Lys	Asp	Thr	
				245					250					255	
Leu	Arg	Lys	Arg	His	Tyr	Trp	Arg	Leu	Asp	Cys	Lys	Cys	Ile	Thr	
				260					265					270	

PF-0683 PCT

Leu	Phe	Gln	Asn	Asn	Thr	Thr	Asn	Arg	Tyr	Tyr	Lys	Glu	Ile	Pro
				275					280					285
Leu	Ser	Glu	Ile	Leu	Thr	Val	Glu	Ser	Ala	Gln	Asn	Phe	Ser	Leu
				290					295					300
Val	Pro	Pro	Gly	Thr	Asn	Pro	His	Cys	Phe	Glu	Ile	Val	Thr	Ala
				305					310					315
Asn	Ala	Thr	Tyr	Phe	Val	Gly	Glu	Met	Pro	Gly	Gly	Thr	Pro	Gly
				320					325					330
Gly	Pro	Ser	Gly	Gln	Gly	Ala	Glu	Ala	Ala	Arg	Gly	Trp	Glu	Thr
				335					340					345
Ala	Ile	Arg	Gln	Ala	Leu	Met	Pro	Val	Ile	Leu	Gln	Asp	Ala	Pro
				350					355					360
Ser	Ala	Pro	Gly	His	Ala	Pro	His	Arg	Gln	Ala	Ser	Leu	Ser	Ile
				365					370					375
Ser	Val	Ser	Asn	Ser	Gln	Ile	Gln	Glu	Asn	Val	Asp	Ile	Ala	Thr
				380					385					390
Val	Tyr	Gln	Ile	Phe	Pro	Asp	Glu	Val	Leu	Gly	Ser	Gly	Gln	Phe
				395					400					405
Gly	Val	Val	Tyr	Gly	Gly	Lys	His	Arg	Lys	Thr	Gly	Arg	Asp	Val
				410					415					420
Ala	Val	Lys	Val	Ile	Asp	Lys	Leu	Arg	Phe	Pro	Thr	Lys	Gln	Glu
				425					430					435
Ser	Gln	Leu	Arg	Asn	Glu	Val	Ala	Ile	Leu	Gln	Ser	Leu	Arg	His
				440					445					450
Pro	Gly	Ile	Val	Asn	Leu	Glu	Cys	Met	Phe	Glu	Thr	Pro	Glu	Lys
				455					460					465
Val	Phe	Val	Val	Met	Glu	Lys	Leu	His	Gly	Asp	Met	Leu	Glu	Met
				470					475					480
Ile	Leu	Ser	Ser	Glu	Lys	Gly	Arg	Leu	Pro	Glu	Arg	Leu	Thr	Lys
				485					490					495
Phe	Leu	Ile	Thr	Gln	Ile	Leu	Val	Ala	Leu	Arg	His	Leu	His	Phe
				500					505					510
Lys	Asn	Ile	Val	His	Cys	Asp	Leu	Lys	Pro	Glu	Asn	Val	Leu	Leu
				515					520					525
Ala	Ser	Ala	Asp	Pro	Phe	Pro	Gln	Val	Lys	Leu	Cys	Asp	Phe	Gly
				530					535					540
Phe	Ala	Arg	Ile	Ile	Gly	Glu	Lys	Ser	Phe	Arg	Arg	Ser	Val	Val
				545					550					555
Gly	Thr	Pro	Ala	Tyr	Leu	Ala	Pro	Glu	Val	Leu	Leu	Asn	Gln	Gly
				560					565					570
Tyr	Asn	Arg	Ser	Leu	Asp	Met	Trp	Ser	Val	Gly	Val	Ile	Met	Tyr
				575					580					585
Val	Ser	Leu	Ser	Gly	Thr	Phe	Pro	Phe	Asn	Glu	Asp	Glu	Asp	Ile
				590					595					600
Asn	Asp	Gln	Ile	Gln	Asn	Ala	Ala	Phe	Met	Tyr	Pro	Ala	Ser	Pro
				605					610					615
Trp	Ser	His	Ile	Ser	Ala	Gly	Ala	Ile	Asp	Leu	Ile	Asn	Asn	Leu
				620					625					630
Leu	Gln	Val	Lys	Met	Arg	Lys	Arg	Tyr	Ser	Val	Asp	Lys	Ser	Leu
				635					640					645
Ser	His	Pro	Trp	Leu	Gln	Glu	Tyr	Gln	Thr	Trp	Leu	Asp	Leu	Arg
				650					655					660
Glu	Leu	Glu	Gly	Lys	Met	Gly	Glu	Arg	Tyr	Ile	Thr	His	Glu	Ser
				665					670					675
Asp	Asp	Ala	Arg	Trp	Glu	Gln	Phe	Ala	Ala	Glu	His	Pro	Leu	Pro
				680					685					690
Gly	Ser	Gly	Leu	Pro	Thr	Asp	Arg	Asp	Leu	Gly	Gly	Ala	Cys	Pro
				695					700					705
Pro	Gln	Asp	His	Asp	Met	Gln	Gly	Leu	Ala	Glu	Arg	Ile	Ser	Val
				710					715					720
Leu														



```
<210> 8
<211> 249
<212> PRT
<213> Homo sapiens
```

```
<220>  
<221> misc_feature  
<223> Incyte ID No: 1511326CD1
```

<400>	8
Met Ala Ser Ser Asp Glu Asp Gly Thr Asn Gly Gly Ala Ser Glu	15
1 5	10
Ala Gly Glu Asp Arg Glu Ala Pro Gly Gln Arg Arg Arg Leu Gly	30
20 25	35
Phe Leu Ala Thr Ala Trp Leu Thr Phe Tyr Asp Ile Ala Met Thr	45
35 40	45
Ala Gly Trp Leu Val Leu Ala Ile Ala Met Val Arg Phe Tyr Met	60
50 55	60
Glu Lys Gly Thr His Arg Gly Leu Tyr Lys Ser Ile Gln Lys Thr	75
65 70	75
Leu Lys Phe Phe Gln Thr Phe Ala Leu Leu Glu Ile Val His Cys	90
80 85	90
Leu Ile Gly Ile Val Pro Thr Ser Val Ile Val Thr Gly Val Gln	105
95 100	105
Val Ser Ser Arg Ile Phe Met Val Trp Leu Ile Thr His Ser Ile	120
110 115	120
Lys Pro Ile Gln Asn Glu Glu Ser Val Val Leu Phe Leu Val Ala	135
125 130	135
Trp Thr Val Thr Glu Ile Thr Arg Tyr Ser Phe Tyr Thr Phe Ser	150
140 145	150
Leu Leu Asp His Leu Pro Tyr Phe Ile Lys Trp Ala Arg Tyr Asn	165
155 160	165
Phe Phe Ile Ile Leu Tyr Pro Val Gly Val Ala Gly Glu Leu Leu	180
170 175	180
Thr Ile Tyr Ala Ala Leu Pro His Val Lys Lys Thr Gly Met Phe	195
185 190	195
Ser Ile Arg Leu Pro Asn Lys Tyr Asn Val Ser Phe Asp Tyr Tyr	210
200 205	210
Tyr Phe Leu Leu Ile Thr Met Ala Ser Tyr Ile Pro Leu Phe Pro	225
215 220	225
Gln Leu Tyr Phe His Met Leu Arg Gln Arg Arg Lys Val Leu His	240
230 235	240
Gly Glu Val Ile Val Glu Lys Asp Asp	245
245	

```
<210> 9
<211> 146
<212> PRT
<213> Homo sapiens
```

```
<220>  
<221> misc_feature  
<223> Incyte ID No: 1519120CD1
```

```

<400> 9
Met Ala Asp Asp Asp Val Leu Phe Glu Asp Val Tyr Glu Leu Cys
  1          5          10          15
Glu Val Ile Gly Lys Gly Pro Phe Ser Val Val Arg Arg Cys Ile
          20          25          30

```

Asn	Arg	Glu	Thr	Gly	Gln	Gln	Phe	Ala	Val	Lys	Ile	Val	Asp	Val
				35					40					45
Ala	Lys	Phe	Thr	Ser	Ser	Pro	Gly	Leu	Ser	Thr	Glu	Gly	Lys	Arg
				50					55					60
Trp	Ile	Ser	Asn	Leu	Lys	Arg	Glu	Ala	Ser	Ile	Cys	His	Met	Leu
				65					70					75
Lys	His	Pro	His	Ile	Val	Glu	Leu	Leu	Glu	Thr	Tyr	Ser	Ser	Asp
				80					85					90
Gly	Met	Leu	Tyr	Met	Val	Phe	Glu	Phe	Met	Asp	Gly	Ala	Asp	Leu
				95					100					105
Cys	Phe	Glu	Ile	Val	Lys	Arg	Ala	Asp	Ala	Gly	Phe	Val	Tyr	Ser
				110					115					120
Glu	Ala	Val	Ala	Ser	Ile	Leu	Asp	Lys	His	Ser	Trp	Lys	Gln	Leu
				125					130					135
Gly	Asp	His	Leu	Asn	Thr	Ala	Leu	Ser	Ser	Ala				
				140					145					

```
<220>  
<221> misc_feature  
<223> Incyte ID No: 1673761CD1
```

10/24

PF-0683 PCT

	245		250		255
Gly Gln Tyr Gly	Val Val Tyr Leu Cys	Asp Asn Trp Gly Gly	His		
	260		270		
Phe Pro Cys Ala	Leu Lys Ser Val Val	Pro Pro Asp Glu Lys	His		
	275		285		
Trp Asn Asp Leu	Ala Leu Glu Phe His	Tyr Met Arg Ser Leu	Pro		
	290		300		
Lys His Glu Arg	Leu Val Asp Leu His	Gly Ser Val Ile Asp	Tyr		
	305		315		
Asn Tyr Gly Gly	Gly Ser Ser Ile Ala	Val Leu Leu Ile Met	Glu		
	320		330		
Arg Leu His Arg	Asp Leu Tyr Thr Gly	Leu Lys Ala Gly Leu	Thr		
	335		345		
Leu Glu Thr Arg	Leu Gln Ile Ala Leu	Asp Val Val Glu Gly	Ile		
	350		360		
Arg Phe Leu His	Ser Gln Gly Leu Val	His Arg Asp Ile Lys	Leu		
	365		375		
Lys Asn Val Leu	Leu Asp Lys Gln Asn	Arg Ala Lys Ile Thr	Asp		
	380		390		
Leu Gly Phe Cys	Lys Pro Glu Ala Met	Met Ser Gly Ser Ile	Val		
	395		405		
Gly Thr Pro Ile	His Met Ala Pro Glu	Leu Phe Thr Gly Lys	Tyr		
	410		420		
Asp Asn Ser Val	Asp Val Tyr Ala Phe	Gly Ile Leu Phe Trp	Tyr		
	425		435		
Ile Cys Ser Gly	Ser Val Lys Leu Pro	Glu Ala Phe Glu Arg	Cys		
	440		450		
Ala Ser Lys Asp	His Leu Trp Asn Asn	Val Arg Arg Gly Ala	Arg		
	455		465		
Pro Glu Arg Leu	Pro Val Phe Asp Glu	Glu Cys Trp Gln Leu	Met		
	470		480		
Glu Ala Cys Trp	Asp Gly Asp Pro Leu	Lys Arg Pro Leu Leu	Gly		
	485		495		
Ile Val Gln Pro	Met Leu Gln Gly Ile	Met Asn Arg Leu Cys	Lys		
	500		510		
Ser Asn Ser Glu	Gln Pro Asn Arg Gly	Leu Asp Asp Ser Thr			
	515		520		

&lt;210&gt; 11

&lt;211&gt; 509

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 1270442CD1

&lt;400&gt; 11

Met Arg Leu Arg	Glu Arg Ser Leu Arg	Gln Asp Pro Asp Leu Arg
1	5	10
Gln Glu Leu Ala	Ser Leu Ala Arg Gly	Cys Asp Phe Val Leu Pro
	20	25
Ser Arg Phe Lys	Lys Arg Leu Lys Ala	Phe Gln Gln Val Gln Thr
	35	40
Arg Lys Glu Glu	Pro Leu Pro Pro Ala	Thr Ser Gln Ser Ile Pro
	50	55
Thr Phe Tyr Phe	Pro Arg Gly Arg Pro	Gln Asp Ser Val Asn Val
	65	70
Asp Ala Val Ile	Ser Lys Ile Glu Ser	Thr Phe Ala Arg Phe Pro
	80	85
		90

PF-0683 PCT

His	Glu	Arg	Ala	Thr	Met	Asp	Asp	Met	Gly	Leu	Val	Ala	Lys	Ala	95	100	105
Cys	Gly	Cys	Pro	Leu	Tyr	Trp	Lys	Gly	Pro	Leu	Phe	Tyr	Gly	Ala	110	115	120
Gly	Gly	Glu	Arg	Thr	Gly	Ser	Val	Ser	Val	His	Lys	Phe	Val	Ala	125	130	135
Met	Trp	Arg	Lys	Ile	Leu	Gln	Asn	Cys	His	Asp	Asp	Ala	Ala	Lys	140	145	150
Phe	Val	His	Leu	Leu	Met	Ser	Pro	Gly	Cys	Asn	Tyr	Leu	Val	Gln	155	160	165
Glu	Asp	Phe	Val	Pro	Phe	Leu	Gln	Asp	Val	Val	Asn	Thr	His	Pro	170	175	180
Gly	Leu	Ser	Phe	Leu	Lys	Glu	Ala	Ser	Glu	Phe	His	Ser	Arg	Tyr	185	190	195
Ile	Thr	Thr	Val	Ile	Gln	Arg	Ile	Phe	Tyr	Ala	Val	Asn	Arg	Ser	200	205	210
Trp	Ser	Gly	Arg	Ile	Thr	Cys	Ala	Glu	Leu	Arg	Arg	Ser	Ser	Phe	215	220	225
Leu	Gln	Asn	Val	Ala	Leu	Leu	Glu	Glu	Glu	Ala	Asp	Ile	Asn	Gln	230	235	240
Leu	Thr	Glu	Phe	Phe	Ser	Tyr	Glu	His	Phe	Tyr	Val	Ile	Tyr	Cys	245	250	255
Lys	Phe	Trp	Glu	Leu	Asp	Thr	Asp	His	Asp	Leu	Leu	Ile	Asp	Ala	260	265	270
Asp	Asp	Leu	Ala	Arg	His	Asn	Asp	His	Ala	Leu	Ser	Thr	Lys	Met	275	280	285
Ile	Asp	Arg	Ile	Phe	Ser	Gly	Ala	Val	Thr	Arg	Gly	Arg	Lys	Val	290	295	300
Gln	Lys	Glu	Gly	Lys	Ile	Ser	Tyr	Ala	Asp	Phe	Val	Trp	Phe	Leu	305	310	315
Ile	Ser	Glu	Glu	Asp	Lys	Lys	Thr	Pro	Thr	Ser	Ile	Glu	Tyr	Trp	320	325	330
Phe	Arg	Cys	Met	Asp	Leu	Asp	Gly	Asp	Gly	Ala	Leu	Ser	Met	Phe	335	340	345
Glu	Leu	Glu	Tyr	Phe	Tyr	Glu	Glu	Gln	Cys	Arg	Ser	Val	Asp	Ser	350	355	360
Met	Ala	Ile	Glu	Ala	Leu	Pro	Phe	Gln	Asp	Cys	Leu	Cys	Gln	Met	365	370	375
Leu	Asp	Leu	Val	Lys	Pro	Arg	Thr	Glu	Gly	Lys	Ile	Thr	Leu	Gln	380	385	390
Asp	Leu	Lys	Arg	Cys	Lys	Leu	Ala	Asn	Val	Phe	Phe	Asp	Thr	Phe	395	400	405
Phe	Asn	Ile	Glu	Lys	Tyr	Leu	Asp	His	Glu	Gln	Lys	Glu	Gln	Ile	410	415	420
Ser	Leu	Leu	Arg	Asp	Gly	Asp	Ser	Gly	Gly	Pro	Glu	Leu	Ser	Asp	425	430	435
Trp	Glu	Lys	Tyr	Ala	Ala	Glu	Glu	Tyr	Asp	Ile	Leu	Val	Ala	Glu	440	445	450
Glu	Thr	Ala	Gly	Glu	Pro	Trp	Glu	Asp	Gly	Phe	Glu	Ala	Glu	Leu	455	460	465
Ser	Pro	Val	Glu	Gln	Lys	Leu	Ser	Ala	Leu	Arg	Ser	Pro	Leu	Ala	470	475	480
Gln	Arg	Pro	Phe	Phe	Glu	Ala	Pro	Ser	Pro	Leu	Gly	Ala	Val	Asp	485	490	495
Leu	Tyr	Glu	Tyr	Ala	Cys	Gly	Asp	Glu	Asp	Leu	Glu	Pro	Leu		500	505	

<210> 12  
 <211> 142  
 <212> PRT

PF-0683 PCT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 1877133CD1

&lt;400&gt; 12

Met	Ile	Ser	Thr	Ala	Arg	Val	Pro	Ala	Asp	Lys	Pro	Val	Arg	Ile
1				5					10					15
Ala	Phe	Ser	Leu	Asn	Asp	Ala	Ser	Asp	Asp	Thr	Pro	Pro	Glu	Asp
				20					25					30
Ser	Ile	Pro	Leu	Val	Phe	Pro	Glu	Leu	Asp	Gln	Gln	Leu	Gln	Pro
				35					40					45
Leu	Pro	Pro	Cys	His	Asp	Ser	Glu	Glu	Ser	Met	Glu	Val	Phe	Lys
				50					55					60
Gln	His	Cys	Gln	Ile	Ala	Glu	Glu	Tyr	His	Glu	Val	Lys	Lys	Glu
				65					70					75
Ile	Thr	Leu	Leu	Glu	Gln	Arg	Lys	Lys	Glu	Leu	Ile	Ala	Lys	Leu
				80					85					90
Asp	Gln	Ala	Glu	Lys	Glu	Lys	Val	Asp	Ala	Ala	Glu	Leu	Val	Arg
				95					100					105
Glu	Phe	Glu	Ala	Leu	Thr	Glu	Glu	Asn	Arg	Thr	Leu	Arg	Leu	Ala
				110					115					120
Gln	Ser	Gln	Cys	Val	Glu	Gln	Leu	Glu	Lys	Leu	Arg	Ile	Gln	Tyr
				125					130					135
Gln	Lys	Arg	Gln	Gly	Ser	Ser								
				140										

&lt;210&gt; 13

&lt;211&gt; 221

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 2636759CD1

&lt;400&gt; 13

Met	Thr	Ser	Gly	Glu	Val	Lys	Thr	Ser	Leu	Lys	Asn	Ala	Tyr	Ser
1				5					10					15
Ser	Ala	Lys	Arg	Leu	Ser	Pro	Lys	Met	Glu	Glu	Gly	Glu	Glu	
				20					25					30
Glu	Asp	Tyr	Cys	Thr	Pro	Gly	Ala	Phe	Glu	Leu	Glu	Arg	Leu	Phe
				35					40					45
Trp	Lys	Gly	Ser	Pro	Gln	Tyr	Thr	His	Val	Asn	Glu	Val	Trp	Pro
				50					55					60
Lys	Leu	Tyr	Ile	Gly	Asp	Glu	Ala	Thr	Ala	Leu	Asp	Arg	Tyr	Arg
				65					70					75
Leu	Gln	Lys	Ala	Gly	Phe	Thr	His	Val	Leu	Asn	Ala	Ala	His	Gly
				80					85					90
Arg	Trp	Asn	Val	Asp	Thr	Gly	Pro	Arg	Leu	Leu	Pro	Arg	His	Gly
				95					100					105
His	Pro	Val	Pro	Arg	Arg	Gly	Gly	Pro	Thr	Thr	Cys	Pro	Pro	Phe
				110					115					120
Asp	Leu	Ser	Val	Phe	Phe	Tyr	Pro	Ala	Ala	Ala	Phe	Ile	Asp	Arg
				125					130					135
Ala	Leu	Ser	Asp	Asp	His	Ser	Lys	Ile	Leu	Val	His	Cys	Val	Met
				140					145					150
Gly	Arg	Ser	Arg	Ser	Ala	Thr	Leu	Val	Leu	Ala	Tyr	Leu	Met	Ile
				155					160					165

PF-0683 PCT

His	Lys	Asp	Met	Thr	Leu	Val	Asp	Ala	Ile	Gln	Gln	Val	Ala	Lys
				170					175					180
Asn	Arg	Cys	Val	Leu	Pro	Asn	Arg	Gly	Phe	Leu	Lys	Gln	Leu	Arg
				185					190					195
Glu	Leu	Asp	Lys	Gln	Leu	Val	Gln	Gln	Arg	Arg	Arg	Ser	Gln	Arg
				200					205					210
Gln	Asp	Gly	Glu	Glu	Glu	Asp	Asp	Arg	Glu	Leu				
				215					220					

&lt;210&gt; 14

&lt;211&gt; 462

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 2716815CD1

&lt;400&gt; 14

Met	Ser	Ile	Ala	Gly	Val	Ala	Ala	Gln	Glu	Ile	Arg	Val	Pro	Leu
1				5					10					15
Lys	Thr	Gly	Phe	Leu	His	Asn	Gly	Arg	Ala	Met	Gly	Asn	Met	Arg
				20					25					30
Lys	Thr	Tyr	Trp	Ser	Ser	Arg	Ser	Glu	Phe	Lys	Asn	Asn	Phe	Leu
				35					40					45
Asn	Ile	Asp	Pro	Ile	Thr	Met	Ala	Tyr	Ser	Leu	Asn	Ser	Ser	Ala
				50					55					60
Gln	Glu	Arg	Leu	Ile	Pro	Leu	Gly	His	Ala	Ser	Lys	Ser	Ala	Pro
				65					70					75
Met	Asn	Gly	His	Cys	Phe	Ala	Glu	Asn	Gly	Pro	Ser	Gln	Lys	Ser
				80					85					90
Ser	Leu	Pro	Pro	Leu	Leu	Ile	Pro	Pro	Ser	Glu	Asn	Leu	Gly	Pro
				95					100					105
His	Glu	Glu	Asp	Gln	Val	Val	Cys	Gly	Phe	Lys	Lys	Leu	Thr	Val
				110					115					120
Asn	Gly	Val	Cys	Ala	Ser	Thr	Pro	Pro	Leu	Thr	Pro	Ile	Lys	Asn
				125					130					135
Ser	Pro	Ser	Leu	Phe	Pro	Cys	Ala	Pro	Leu	Cys	Glu	Arg	Gly	Ser
				140					145					150
Arg	Pro	Leu	Pro	Pro	Leu	Pro	Ile	Ser	Glu	Ala	Leu	Ser	Leu	Asp
				155					160					165
Asp	Thr	Asp	Cys	Glu	Val	Glu	Phe	Leu	Thr	Ser	Ser	Asp	Thr	Asp
				170					175					180
Phe	Leu	Leu	Glu	Asp	Ser	Thr	Leu	Ser	Asp	Phe	Lys	Tyr	Asp	Val
				185					190					195
Pro	Gly	Arg	Arg	Ser	Phe	Arg	Gly	Cys	Gly	Gln	Ile	Asn	Tyr	Ala
				200					205					210
Tyr	Phe	Asp	Thr	Pro	Ala	Val	Ser	Ala	Ala	Asp	Leu	Ser	Tyr	Val
				215					220					225
Ser	Asp	Gln	Asn	Gly	Gly	Val	Pro	Asp	Pro	Asn	Pro	Pro	Pro	Pro
				230					235					240
Gln	Thr	His	Arg	Arg	Leu	Arg	Arg	Ser	His	Ser	Gly	Pro	Ala	Gly
				245					250					255
Ser	Phe	Asn	Lys	Pro	Ala	Ile	Arg	Ile	Ser	Asn	Cys	Cys	Ile	His
				260					265					270
Arg	Ala	Ser	Pro	Asn	Ser	Asp	Glu	Asp	Lys	Pro	Glu	Val	Pro	Pro
				275					280					285
Arg	Val	Pro	Ile	Pro	Pro	Arg	Pro	Val	Lys	Pro	Asp	Tyr	Arg	Arg
				290					295					300
Trp	Ser	Ala	Glu	Val	Thr	Ser	Ser	Thr	Tyr	Ser	Asp	Glu	Asp	Arg

PF-0683 PCT

Pro	Pro	Lys	Val	305	Pro	Pro	Arg	Glu	Pro	310	Leu	Ser	Pro	Ser	Asn	Ser	315
				320						325							330
Arg	Thr	Pro	Ser	335	Pro	Lys	Ser	Leu	Pro	340	Ser	Tyr	Leu	Asn	Gly	Val	345
Met	Pro	Pro	Thr	350	Gln	Ser	Phe	Ala	Pro	355	Asp	Pro	Lys	Tyr	Val	Ser	360
				365						370							375
Ser	Lys	Ala	Leu	380	Gln	Arg	Gln	Asn	Ser	385	Glu	Gly	Ser	Ala	Ser	Lys	390
Val	Pro	Cys	Ile	395	Leu	Pro	Ile	Ile	Glu	400	Asn	Gly	Lys	Lys	Val	Ser	405
				410						415							420
Ser	Thr	His	Tyr	425	Tyr	Leu	Leu	Pro	Glu	430	Arg	Pro	Pro	Tyr	Leu	Asp	435
Lys	Tyr	Glu	Lys	440	Phe	Phe	Arg	Glu	Ala	445	Glu	Glu	Thr	Asn	Gly	Gly	450
Ala	Gln	Ile	Gln	455	Pro	Leu	Pro	Ala	Asp	460	Cys	Gly	Ile	Ser	Ser	Ala	
Thr	Glu	Lys	Pro		Asp	Ser	Lys	Thr	Lys		Met	Asp	Leu	Gly	Gly	His	
Val	Lys	Arg	Lys		His	Leu	Ser	Tyr	Val		Val	Ser	Pro				

&lt;210&gt; 15

&lt;211&gt; 2192

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 480457CB1

&lt;400&gt; 15

```

gatcaatgaa gccgagtga tgggggctga atgtgcgagt ccatagctga agaggagcgc 60
cagatgggtgg aggaatacac ttatttatga aactgtcttg agttcttctt gaattgccag 120
ttttcagcct cctcatgcct ccgtctcctt tagacgacag ggtagtagtg gcactatcta 180
ggccccgtccg acctcaggat ctcaaccttt gtttagactc tagttacctt ggctctgcca 240
acccaggcag taacagccac cctcctgtca tcgccaccac cgttgtgtcc ctcaaggctg 300
cgaatctgac gtatatgccc tcatccagcg gctctgcccg ctgctgaat tgtggatgca 360
gcagtgccag ctgctgcaat gtggcaacct acgacaagga caatcaggcc caaacccaag 420
ccattgccgc tggcaccacc accactgcca tcggaacctc taccacctgc cctgctaacc 480
agatgggtcaa caataatgag aatacaggct ctctaagtcc atcaagtggg gtgggcagcc 540
ctgtgtcagg gacccccaaag cagctagcca gcatcaaaat aatctacccc aatgacttgg 600
caaagaagat gaccaaattgc agcaagagtc acctgccgag tcaggggccct gtcattcattg 660
actgcaggcc cttcatggag tacaacaaga gtcacatcca aggagctgtc cacattaact 720
gtgccgataa gatcagccgg cggagactgc agcagggcaa gatcactgtc ctagacttga 780
tttctctgtg ggaaggcaag gactctttca agaggatctt ttccaaagaa attatagttt 840
atgatgagaa taccaatgag ccaagccgag tgatgccttc ccagccactt ccatagttcc 900
tcgagtcctt gaagagagaa ggcaaagaac ctctgtgtgt gaaagggtga cttagtagtt 960
ttaagcagaa ccatgaaaac ctctgtgaca actccttcca gctccaagag tgccgggagg 1020
tggggggcgg cgcattccgg gcctcgagct tgctacctca gcccatcccc accacccttg 1080
acatcgagaa cgctgagctc acccccatct tgcccttccct gttccttggc aatgagcagg 1140
atgctcagga cctggacacc atgcagcggc tgaacatcgg ctacgtcatc aacgtcacca 1200
ctcatcttcc cctctaccac tatgagaaag gacctgttcaa ctacaagcgg ctgccagcca 1260
ctgacagcaa caagcagaac ctgcggcagt accttgaaga ggcttttgag ttcatagagg 1320
aagctcacca gtgtgggaag gggcttctca tccactgcca ggctgggggtg tcccgtctcg 1380
ccaccatcgt catcgcttac ttgatgaagc acactcggat gaccatgact gatgcttata 1440
aatttgtcaa aggcacaacga ccaattatct ccccaaacct taacttcatg gggcagttgc 1500
tagagttcga ggaagacctt aacaacgggt tgacaccgag aatccttaca ccaaagctga 1560
tgggcgtgga gacgggttggt tgacaatggt ctggatggaa aggattgctg ctctccatta 1620
ggagacaatg aggaaggagg atggattctg gttttttttt tttctttttt tttttgtagt 1680

```

PF-0683 PCT

```

tgggagtaag tttgtgaatg gaaacaaact tgtttaaaca ctttattttt aacaagtgtg 1740
agaagactat aactttttgat gccattgaga ttcacctccc acaaactgac aaattaagga 1800
gggttaaagaa gtaatttttt taagccaaca ataaaaatat aatacaactt gtttctcccc 1860
cttttccttt taagctattt gtagagttaa tgactaaata gtctgtgcag gttcatagac 1920
cgaagatact acacacttta aaccaattaa aaagaaccaa aagtaaatag aaaagacatt 1980
gaatcaccaa ggcctgggat caacctgggc tgtccacaca gaaaacaaaa acccaaccaa 2040
accaagccct gttgtgctca ctgggtgcaa gagaagatca gggcagctta agtgggtctaa 2100
gaatcccttca ggcattcttt aaggagaaaa aggatacctt tgattttgtg tgtttcatgc 2160
tctggatttt tttttttttc cttctctggg tt 2192

```

&lt;210&gt; 16

&lt;211&gt; 1338

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; unsure

&lt;222&gt; 959, 1029, 1159

&lt;223&gt; a or g or c or t, unknown, or other

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 563663CB1

&lt;400&gt; 16

```

ccttgtcttg aaaagtaaaa gtggatcctg ccacgttcgg agctccctgg cgcctcgccc 60
ggctggagct agagaactcg tcctgtggcg gccccggcg tggggcgga cagcgcccc 120
ctggaggggg cagtcccggg agaacctgcg gcggcggag cggtaaaaat aagtactaa 180
agaagcagac ctgggaatca cctaactgt cgaggaggag atttgattgc cgaagtattt 240
caggcctact aactacaact cctcaaattc caataaaaat ggaaaacttt aataatttct 300
atatacttac atctaaagag ctaggaggag gaaaatttgc tgtgggttaga caatgtatat 360
caaatctac tggccaagaa tatgctgcaa aattttctaa aaagagaaga agaggacagg 420
attgtcgggc agaaatttta cagcagattg ctgtgcttga attggcaaag tcttgtcccc 480
gtgttattaa tcttcatgag gtctatgaaa atacaagtga aatcattttg atattggaat 540
atgctgcagg tggagaaatt ttcagcctgt gtttacctga gttggctgaa atggtttctg 600
aaaatgatgt tatcagactc attaaacaaa tacttgaagg agtttattat ctacatcaga 660
ataacattgt acaccttgat ttaaagccac agaatatatt actgagcagc atataccctc 720
tcggggacat taaaatagta gatgggggaa tgtctcgaaa aatagggcaa tgcgtgtgac 780
cttcgggaaa tcatgggaac accagaatat ttagctccag aaatcctgaa ctatgatccc 840
attaccacag aaacagatat gtggaatatt ggtataatag catatatgtc gctaactcac 900
acatcaccat ttgtgggaga agataatcaa gaaacatacc tcaatctctc tcaagttant 960
gtagattatt cggaagaaac tttcccatca gtttcacagc tggccacaga ctttattcag 1020
agcccttcng taacacctcc agagaaacga caccagcaga gatctgcctc cctcattctc 1080
gctacagca gcgggacttt gcaacattgt ctcacctgaa cgaacctcca gttccccctc 1140
aactcaggat caatctgtna ggtctctaag acaagactct taatccccctc gcatggaact 1200
cgggtgaaga gaagtacgag atctcccgcg gttgcagact gctcaaacca ttcgtctgat 1260
atctatcccc cccccccacc tgtccaattg gttttacatc tttcttgcctc aggccttctc 1320
gacttgcttc ccccttac 1338

```

&lt;210&gt; 17

&lt;211&gt; 1706

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 1425842CB1

&lt;400&gt; 17



PF-0683 PCT

```

cgtgggagcc ggcggggccg ctgtcgtccc aacccccgcc gccctcgtcg cgcgcggggc 60
ctccgcgccc ccggctgctg ctcacgcccc gcccgggagc cagattttgt ggaagtataa 120
tactttgtca ttatgagatg tcgtctctcg gtgcctcctt tgtgcaaatt aaatttgatg 180
acttgagttt ttttgaaaac tgcggtggag gaagtttttg gagtgtttat cgagccaaat 240
ggatatcaca ggacaaggag gtggctgtaa agaagctcct caaaatagag aaagaggcag 300
aaatactcag tgcctcagc cacagaaaca tcatccagtt ttatggagta attcctgaac 360
ctcccaacta tggcattgtc acagaatatg cttctctggg atcactctat gattacatta 420
acagtaacag aagtgaggag atggatatgg atcacattat gacctgggac actgatgtag 480
ccaaaggaat gcattattta catatggagg ctctgtcaa ggtgattcac agagacctca 540
agtcaagaaa cgttgttata gctgctgatg gagtattgaa gatctgtgac tttgggtgcct 600
ctcgggttcca taaccatata acacacatgt ccttgggttg aactttccca tggatggctc 660
cagaagttat ccagagtctc cctgtgtcag aaacttgtga cacatatcc tatggtgtgg 720
ttctctggga gatgctaaca agggagggtcc cttttaaaagg tttggaagga ttacaagtag 780
cttggcttgt agtggaaaaa aacgagagat taaccattcc aagcagttgc ccagaaagtt 840
ttgctgaact gttacatcag tgttgggaag ctgatgccaa gaaacggcca tcattcaagc 900
aaatcatttc aatcctggag tccatgtcaa atgacacgag ccttctctgac aagtgttaact 960
cattcctaca caacaaggcg gagtggaggt gcgaaattga ggcaactctt gagaggctaa 1020
agaaactaga gcgtgatctc agctttaagg agcaggagct taaagaacga gaaagacgtt 1080
taaagatgtg ggagcaaaaag ctgacagagc agtccaacac cccgcttctc ttgcctcttg 1140
ctgcaagaat gtctgaggag tcttactttg aatctaaaac agaggagtca aacagtgcag 1200
agatgtcatg tcagatcaca gcaacaagta acggggaggg ccatggcatg aacccaagtc 1260
tgagggccat gatgctgatg ggctttgggg atatcttctc aatgaacaaa gcaggagctg 1320
tgatgcattc tgggatgcag ataaacatgc aagccaagca gaattcttcc aaaaccacat 1380
ctaagagaag ggggaagaaa gtcaacatgg ctctgggggt cagtgatattt gacttgtcag 1440
aaggtgacga tgatgatgat gatgacggtg aggaggagga taatgacatg gataatagt 1500
aatgaaagca gaaagcaaaag taataaaatc acaaattgtt ggaaaacaca aaagtaact 1560
gtttatctca gtctgtacaa aaacagtaag gaggcagaaa gccaaagcact gcatttttag 1620
ggcaatcaca tttacatgac cgtatttctt atcattctac tttattctgc tacagaaaac 1680
gcgggagatt agccaagagt tattat 1706

```

&lt;210&gt; 18

&lt;211&gt; 2140

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 2349047CB1

&lt;400&gt; 18

```

aaaaagggtg acaaccggga aagtagactc accgtcttgg tctagagact gacccttgca 60
cagacagacc ccttcccttc tctgcgaaag gaccaagccc cagaagtcac tccatctcct 120
acggctcgca atttccagag gccccctggc acctccagc ctgatgtcgt gcgtcaagtt 180
atggcccagc ggtgcccccg ccccttgggt gtccatcgag gaactggaga accaggagct 240
cgtcggcaaa ggcggttctg gcacagtgtt ccgggcgcaa cataggaagt ggggctacga 300
tgtggcggtc aagatcgtaa actcgtgagt gaccccggtt gacccagcc gcaatctggc 360
caaaaagggtg gagccactgt gctctggggc ctgaatggcg aaggaggga ttttccacg 420
accccggtgc gactccagct ctctcctgga gtgtaggaa gcatatcca gggagggtcaa 480
ggccatggca agtctggata acgaattcgt gctgcgccta gaagggttta tcgagaaggt 540
gaactgggac caagatccca agccggctct ggtgactaaa ttcattggaga acggctcctt 600
gtcggggctg ctgcagtccc agtgccctcg gccctggcgg ctcttttgcc gcctgtgtaa 660
agaagtgggtg cttgggatgt tttacctgca cgaccagaac ccggtgctcc tgcaccggga 720
cctcaagcca tccaacgtcc tgctggaccc agagctgcac gtcaagctgg cagattttgg 780
cctgtccaca tttcagggag gctcacagt agggacagg tccggggagc cagggggcac 840
cttgggtctac ttggcccag aactgtttgt taacgtaaac cggaaggcct ccacagccag 900
tgacgtctac agcttcggga tcctaattgt ggcatgtctt gctggaagag aagttgagtt 960
gccaaaccgaa ccatcactcg tgtacgaagc agtgtgcaac aggcagaacc ggccttcatt 1020
ggctgagctg cccaagccg ggcctgagac tcccggctta gaaggactga aggagctaat 1080
gcagctctgc tggagcagtg agcccaagga cagaccctcc ttccaggaat gcctacaaa 1140
aactgatgaa gtcttccaga tgggtggagaa caatatgaat gctgctgtct ccacggtaaa 1200

```

PF-0683 PCT

```

ggatttcctg tctcagctca ggagcagcaa taggagattt tctatcccag agtcaggcca 1260
aggaggggaca gaaatggatg gcttttaggag aaccatagaa aaccagcact ctcgtaatga 1320
tgtcatgggtt tctgagtggtc taaacaaact gaatctagag gagcctccca gctctgttcc 1380
taaaaaatgc ccgagcctta ccaagaggag cagggcacaa gaggagcagg ttccacaagc 1440
ctggacagca ggcacatctt cagattcgat ggcccaacct cccagactc cagagacctc 1500
aactttcaga aaccagatgc ccagccctac ctcaactgga acaccaagtc ctggaccctc 1560
aggggaatcag ggggctgaga gacaaggcat gaactggtcc tgcaggaccc cggagccaaa 1620
tccagtaaca gggcgaccgc tcgttaacat atacaactgc tctgggggtgc aagttggaga 1680
caacaactac ttgactatgc aacagacaac tgccttgccc acatggggct tggcaccttc 1740
gggcaagggg aggggcttgc agcaccccc accagtaggt tcgcaagaag gccctaaaga 1800
tccatgaagcc tggagcaggc cacagggttg gtataatcat agcgggaaat aaagcacctt 1860
ccaagcttgc ctccaagagt tacgagttaa ggaagagtgc cacccttga ggcccctgac 1920
ttccttctag ggcagtctgg cctgcccaca aactgacttt gtgacctgtc cccagcaggt 1980
caataaacat gatggaaata aagcaccttc caagcttgcc tccaagagtt acgagttaag 2040
gaagagtgcc accccttgag gcccctgact tccttctagg gcagtctggc ctgcccacaa 2100
actgactttg tgacctgtcc cccaggagtc aataaacatg
2140

```

&lt;210&gt; 19

&lt;211&gt; 1573

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 2415617CB1

&lt;400&gt; 19

```

gccccacagc cggccctgcg acgcccgcct gggcagcacc gataaggagc tgaaggcagg 60
agccgcgcgc acgggcagcg cccacagcg ccagggaccc cctggcagcg ggagccgcgg 120
gtcgagggtta tggatccagc gggcgcccc cggggcggtg tcccgcggcc ctgccgcgtg 180
ctggtgctgc tgaaccgcgc cggcggaag ggcaaggcct tgcagctctt ccggagtcac 240
gtgcagcccc ttttggctga ggctgaaatc tccttcacgc tgatgctcac tgagcggcgg 300
aaccacgcgc gggagctggt gcggtcggag gagctgggcc gctgggacgc tctggtggtc 360
atgtctggag acgggctgat gcacgaggtg gtgaacgggc tcatggagcg gcctgactgg 420
gagaccgcca tccagaagcc cctgtgtagc ctcccagcag gctctggcaa cgcgctggca 480
gcttccttga accattatgc tggctatgag caggtcacca atgaagacct cctgaccaac 540
tgcacgctat tgctgtgccg ccggtgctg tcacccatga acctgctgtc tctgcacacg 600
gcttcggggc tgcgcctctt ctctgtgctc agcctggcct ggggcttcat tgctgatgtg 660
gacctagaga gtgagaagta tcggcgtctg ggggagatgc gcttcactct gggcaccttc 720
ctcgctctgg cagccctgcg cacctaccgc ggccgactgg cctacctccc tgtaggaaga 780
gtgggttcca agacacctgc ctcccccggt gtggtccagc agggcccggt agatgcacac 840
cttgtgccac tggaggagcc agtgccctct cactggacag tggtgcccga cgaggacttt 900
gtgctagtc tggcactgct gcactcgac ctgggcagtg agatgtttgc tgcacccatg 960
ggcgcgtgtg cagctggcgt catgcatctg ttctacgtgc gggcgggagt gtctcgtgcc 1020
atgtctgtgc gcctcttctt ggccatggag aagggcaggc atatggagta tgaatgcccc 1080
tacttggtat atgtgcccgt ggtcgccctc cgcttgagc ccaaggatgg gaaaggtgtg 1140
tttgcaagtgg atggggaatt gatggttagc gagggcgtgc agggccaggt gcacccaaac 1200
tacttctgga tggtcagcgg ttgcgtggag cccccgcca gctggaagcc ccagcagatg 1260
ccaccgccag aagagccctt atgacccttg ggccgcgctg tgccttagtg tctacttgca 1320
ggacccttcc tccttcctta gggctgcagg gcctgtccac agctcctgtg ggggtggagg 1380
agactcctct ggagaagggt gagaagggtg aggcctatgt ttggggggac aggccagaat 1440
gaagtccttg gtcaggagcc cagctggctg ggcccagctg cctatgtaag gccttctagt 1500
ttgttctgag acccccaccc cacgaaccaa atccaaataa agtgacattc ccagcctgaa 1560
aaaaaaaaaa aaa
1573

```

&lt;210&gt; 20

&lt;211&gt; 862

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

PF-0683 PCT

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 3815186CB1

&lt;400&gt; 20

```

aaccacagaga gtaggggtggt catgattctg tgggtgggtgca ggcattgtttt ctttagcatta 60
tttaccatca tattacgtat aagaaagctg acagtggaga gagattaaaa gcaggactat 120
taggggatag agagtaattg gaggcctttt tataatgtgc agattcgtga tttttgaaaa 180
gcatagcttt gattctgctt tccttttttt aacttggttt tttaaaattt gcacacacac 240
accactttat ggcaattctt aactgacatt caatgactta cttcttttct tagaaaattt 300
ccaccacatt tctatcccca agccaacata caatgtgaaa tgaaagccag tgcgtggagt 360
gcagctgcta aaaattttca gcacagggtt ctttctgact ctgctcatga gatggtatca 420
gccacccaat gactggcgta tcttggtcct gtgtccttct tcttacgctg tgtaaatgtg 480
tttacttttc atttggcaga gagacaagag agacacctcc aacttcgaca aagagttcac 540
cagacagcct gtggaactga ccccaactga taaactcttc atcatgaact tggacaaaaa 600
tgaatttgct ggcttctcgt atactaacc agagtttgtc attaatgtgt aggtgaatgc 660
agattccatc gctgagcctg tgtgtaaggc tgcaggctga atgtctatta tcaattccag 720
tcttccagga tcatggtgct ctctgttggc atccgtcatg tggagagctt gtcttagagg 780
gcttttcttt gtatgtatag cttgctagtt tgttttctac atttcaaaat gtttagttta 840
gaataagtgc attgccact ga 862

```

&lt;210&gt; 21

&lt;211&gt; 2744

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 5504544CB1

&lt;400&gt; 21

```

ccggcccccgg gttccgggggt ctcttttcac atccagatcg ggctgaccgg cgagttcgtg 60
ctgttgcccg ccgcctccga gctggctcat gtgaagcagc tggcctgttc catcgtggac 120
cagaagttcc ctgagtgtgg cttctacggc ctttacgaca agatcctgct tttcaaact 180
gacccacgt cgcccaacct cctgcagctg gtgcgctcgt ccggagacat ccaggagggc 240
gacctggtg aggtgggtgct gtcggcctcg gccaccttcg aggacttcca gatccgccc 300
cacgacctca cgggtgcactc ctatcggggc cctgccttct gtgatcactg cggggagatg 360
ctcttcggcc tagtgcgcca gggcctcaag tgcgatggct gcgggctgaa ctaccacaag 420
cgctgtgcct tcagcatccc caacaactgt agtggggccc gcaaaccggc cctgtcatcc 480
acgtctctgg ccagtggcca ctcggtgcgc ctccggcacct ccgagtcctt gccctgcacg 540
gctgaagagc tgagccgtag caccaccgaa ctctcgctc gccgtccccc gtcactctct 600
tctcctctt ctgcctcatc gtatacgggc cgccccattg agctggacaa gatgctgctc 660
tccaagggtc aggtgccgca cacttctct atccacagct atacacggcc caccgtttgc 720
caggcttgca agaaactcct caagggcctc ttccggcagg gcctgcaatg caaagactgc 780
aagtttaact gtcacaaacg ctgcgccacc cgcgtcccta atgactgcct gggggaggcc 840
cttatcaatg gagatgtgcc gatggaggag gccaccgatt tcagcgaggc tgacaagagc 900
gccctcatgg atgagtcaga ggactccggt gtcactccct gctccactc agagaatgcg 960
ctccacgcca gtgaggagga ggaaggcgag ggaggcaagg cccagagctc cctgggggtac 1020
atccccctaa tgagggtggt gcaatcgggt cgacacacga cgcggaatc cagcaccacg 1080
ctgcgggagg gttgggtggt tcattacagc aacaaggaca cgctgagaaa gcggcactat 1140
tggcgccctg actgcaagt tatcacgctc ttccagaaca acacgaccaa cagatactat 1200
aaggaaattc cgctgtcaga aatcctcagc gtggagtccg ccagaaactt cagccttggt 1260
ccgcccgggc ccaaccaca ctgctttgag atcgctactg ccaatgccac ctacttcgtg 1320
ggcgagatgc ctggcgggac tccgggtggg ccaagtgggc agggggctga ggccgcccgg 1380
ggctgggaga cagccatccg ccaggccctg atgcccgtca tccttcagga cgcaccacg 1440
gccccaggcc acgcgccccca cagacaagct tctctgagca tctctgtgtc caacagtcag 1500
atccaagaga atgtggacat tgccactgtc taccagatct tccctgacga agtgcctggg 1560
tcagggcagt ttggagtggg ctatggagga aaacaccgga agacaggccg ggacgtggca 1620
gttaagggtc ttgacaaact gcgcttccct accaagcagg agagccagct ccggaatgaa 1680
gtggccattc tgcagagcct gcggcatccc gggatcgtga acctggagtg catgttcgag 1740

```

PF-0683 PCT

```

acgcctgaga aagtgtttgt ggtgatggag aagctgcatg gggacatggt ggagatgata 1800
ctgtccagtg agaagggccg gctgcctgag cgcctcacca agttcctcat caccagatc 1860
ctggtggctt tgagacacct tcaattcaag aacattgtcc actgtgactt gaaaccagaa 1920
aacgtgttgc tggcatcagc agaccattt cctcaggtga agctgtgtga ctttggcttt 1980
gctcgcatca tcggcgagaa gtcgttcgc cgctcagtgg tgggcacgcc ggcctacctg 2040
gcacccgagg tgctgtctca ccagggtac aaccgctcgc tggacatgtg gtcagtgggc 2100
gtgatcatgt acgtcagcct cagcggcacc ttccttttca acgaggatga ggacatcaat 2160
gaccagatcc agaacgccgc cttcatgtac cgggccagcc cctggagcca catctcagct 2220
ggagccattg acctcatcaa caacctgctg caggtgaaga tgcgcaaacg ctacagcgtg 2280
gacaaatctc tcagccaccc ctggttacag gagtaccaga cgtggctgga cctccgagag 2340
ctggagggga agatgggaga gcgatacatc acgcattgaga gtgacgacgc gcgctgggag 2400
cagtttgcag cagagcatcc gctgcctggg tctgggctgc ccacggacag gcatctcggg 2460
ggggcctgtc caccacagga ccacgacatg caggggctgg cggagcgcag cagtgttctc 2520
tgaggtcctg tgccctcgtc cagctgctgc cctccacagc ggttcttcac aggatccag 2580
caatgaactg ttctagggaa agtggcttcc tgccaaact ggatgggaca cgtggggagt 2640
ggggtggggg gagctatttc caaggccctt ccctgtttcc ccagcaatta aaacggactc 2700
atctctggcc ccatggcctt gatctcagca aaaaaaaaaa aaaa 2744

```

&lt;210&gt; 22

&lt;211&gt; 1266

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 1511326CB1

&lt;400&gt; 22

```

ggggcgctga cggaagcggc ggcagcgggc agcggctctc gggctgcagg ctgggcaggg 60
tccctcccca cgctcctgcc gctgtctccc acgtcccca ggtgcgcggc caccatggcg 120
tccagcgacg aggacggcac caacggcggc gcctcggagg ccggcgagga ccgggaggct 180
cccggccagc ggaggcgctt ggggttcttg gccaccgctt ggctcacctt ctacgacatc 240
gccatgaccg cggggtgggtt ggttctagct attgccatgg tacgttttta tatggaaaaa 300
ggaacacaca gaggtttata taaaagtatt cagaagacac tttaaatttt ccagacattt 360
gccttgcttg agatagttca ctgtttaatt ggaattgtac ctacttctgt gattgtgact 420
ggggtccaag tgagtccaag aatctttatg gtgtggctca ttactcacag tataaaacca 480
atccagaatg aagagagtggt ggtgcttttt ctggctgcgt ggactgtgac agagatcact 540
cgctattcct tctacacatt cagccttctt gaccacttgc catacttcat taaatggggc 600
agatataatt tttttatcat cttatatcct gttggagttg ctggtgaact tcttacaata 660
tacgctgcct tgccgcatgt gaagaaaaca ggaatgtttt caataagact tctaacaaa 720
tacaatgtct cttttgacta ctattatttt cttcttataa ccatggcatc atatatacct 780
ttgtttccac aactctatct tcatatgtta cgtcaaagaa gaaaggtgct tcatggagag 840
gtgattgtag aaaaggatga ttaaattgat tctgcaacaa aggtgctttt tccagaataa 900
ccaagattac ctgagtccaa gttttaataa caagaataaa caactttgtg aaatatcatg 960
gattgtatgg tttcttaaaa tataacttga gacacgtggt atttgccagt atttgtgttc 1020
ctcttgtgcc agatctatct tttacaagaa ctgtgcaaat atcagtaact tttgggtagg 1080
tattgattat taggaaaata attaggtgta ttatctgggg gaaaaaaaaa cttttgctaa 1140
gttttttttg aaacatgctc aaagcttttt aaatcaatat ttagaaatta gtttaacgat 1200
ttactattat acctgctagt gatatttatg tgatatttat aaatgaaaaa aaatgcaaaa 1260
ttatat 1266

```

&lt;210&gt; 23

&lt;211&gt; 594

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 1519120CB1

PF-0683 PCT

&lt;400&gt; 23

```

cgggccgcta tccccctccgg accatggccg acgacgacgt gctgttcgag gatgtgtacg 60
agctgtgcga ggtgatcgga aagggtccct tcagtgttgt acgacgatgt atcaacagag 120
aaactgggca acaatttgct gtaaaaattg ttgatgtagc caagttcaca tcaagtccag 180
ggttaagtac agaaggtaag agatggattt caaatctaaa gcgggaagcc agtatctgtc 240
atatgttgaa acatccacac attgttagagt tattggagac atatagctca gatggaatgc 300
tttacatggt ttttgaattt atggatggag cagatctgtg ttttgaaatc gtaaagcgag 360
ctgacgctgg ttttgtgtac agtgaagctg tagccagcat cttagacaaa catagctgga 420
aacaacttgg agaccacctt aatacagctt tgagctcagc atgacttgaa tgtcttcaaa 480
ttcattgccc tgactccttc tggctgaaga gcaagggacc tagtctcagg gatggagcag 540
ccatggctat ggcacaaaat aaagaatatg ctcaatgaaa aaaaaaaaaa aaaa 594

```

&lt;210&gt; 24

&lt;211&gt; 2280

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 1673761CB1

&lt;400&gt; 24

```

cccaaacgctc tggaatatac tcgaaaaaag gagaatgagt tgtatgaatc attgatgaat 60
attgccaacc gaaagcagga ggaaatgaag gatatgattg ttgagacact taataccatg 120
aaggaggaac ttctggatga tgctactaac atggagttaa aagacgtcat tgtccctgag 180
aatggagaac cagtaggcac cagagagatc aaatgctgca tccgacagat ccaggaaactc 240
atcatctccc gacttaatac ggacgtggct aataagctga tcagctcagt ggattacctg 300
agggaaaagct tcgtcggaac cctggaacga tgtctgcaga gcctggagaa gtctcaggat 360
gtctcagttc acatcaccag taattatctc aaacagatct taaatgctgc ctatcatgtt 420
gaagtcacgt ttactcagg gtcgtcagtt acaaggatgc tatgggagca aatcaaacag 480
atcatccagc gcatcacatg ggtgagccca cctgccatca ctctggaatg gaagaggaag 540
gtggcccagg aagccattga gagcctcagc gcctccaaat tggctaagag catttgacgc 600
caattccgga ctcggtcaa tagttcccac gaggtttttg cagcctcctt gcggcagctg 660
gaagctggcc actcaggccg gttagagaaa acggaagatc tatggctgag ggctcggaaa 720
gatcatgctc cccgcctggc ccgcctttct ctggaaagct gttctttaca ggtgtcttg 780
cttcatcgta aacctaaact gggacaggaa ctgggcccgg gccagtatgg tgtggtatc 840
ctgtgtgaca actggggagg acacttccct tgtgccctca aatcagttgt cctccagat 900
gagaagcact ggaatgatct ggctttggaa tttactata tgagggtctt gccgaagcat 960
gagcgattgg tggatctcca tggttcagtc attgactaca actatggtgg tggctccagc 1020
attgctgtgc tctcattat ggagcggcta caccgggac tctacacagg gctgaagct 1080
gggctgacc tggagacacg tttgcagata gcactagatg tgggtggagg aatccgctt 1140
ctgcacagcc agggacttgt ccactgtgat atcaaaactga aaaatgtgct gctggataag 1200
cagaaccgtg ccaagatcac tgacttagga ttctgcaagc cagaggccat gatgtcaggc 1260
agcattgtgg ggacaccaat ccataatggcc cctgaacttt tcacagggaa gtacgataat 1320
tccgtggatg tctacgcttt tggaaattct ttctggtata tctgctcagg ctctgtcaag 1380
ctccctgagg catttgagag gtgtgctagc aaagaccatc tctggaacaa tgtgcggagg 1440
ggggctcgcc cagaacgtct tcctgtgttt gatgaggagt gctggcagtt gatggaagcc 1500
tggtgggatg gcgacccctt gaagaggcct ctcttgggca ttgtccagcc catgctccag 1560
ggcatcatga atcggtcttg caagtccaat tctgagcagc caaacagagg actagatgat 1620
tctacttgaa agcaaagacc tttctcttct actctctagt tatttccttc cccctcacct 1680
tttggccatg gggagaattt gacatttatt cactatagga cacactccca agggaaactgg 1740
tgcttgctgg gaaacttgga accttcccag gcagggatga ctctgggaca gtgaagagtt 1800
gaatgactga gcatattcag cagctcactg aagcgccaag ctatcccttt agcaaaaaag 1860
tgtctcagat gtgtaaaagc tgaggaaatgt ggtgttcttg cttcacaaat gaaaaggagg 1920
cagatgtttac cattgtcttt tctactgtata tacttctaag acagcaagcg ggacactgca 1980
gtggcaatag tgttaaaaaa tctcattctc atgatttttg gctctagcta ggaatagttt 2040
agtcaggact cagaaaaatt gagcttagtt ctgggattat gaaaacatgg ggacaaaaac 2100
aataacttgt ggatgtctgg ttctgtctgt ctgcagccag gtatctcagg ttgcaatgg 2160
cagaagtccc agtgaggagg ctagaaacag agctatctgt tctcatagt gccagtgtgt 2220
ttacatttat aggaccacat atccctgggt tctgggggag gtttctactgt acacagccaa 2280

```

<400> 26							
aaggaaagga	gagaaggggt	gtgaaggga	gcggaatgga	agggaagggga	ggtcccgtgg		60
gacgctgggg	tctggggcag	agcaggtagc	agcgtgctgc	cctgacagct	gtctccgctc		120
ctcagattgt	cagtggctgc	tatgcagcag	gtgcagcctg	gtctctcact	gagtctctac		180
tccacaaagg	caacgactgg	ccaaggcagt	ggctggctct	gggttacaca	agtgcagaca		240
ctcaactaag	tgagctggaa	gacccaggag	aaggcggagg	ctcaggtgcc	cacatgatca		300
gcacagccag	ggtacctgct	gacaagcctg	tacgcatcgc	ctttagcctc	aatgacgcct		360
catagtatac	accccctgaa	gactccattc	ctttggtcct	tccagaatta	gaccagcagc		420
tacagccctt	gcgccttgtt	catgactccg	aggaatccat	ggaggtgttc	aaacagcact		480
gccaaatagc	agaagaatac	catgaggtca	aaaagaaagt	caccctgctt	gagcaaaagg		540
agaaggagct	cattgccaaag	ttagatcagg	cagaaaagga	gaaggtggat	gctgctgaac		600

PF-0683 PCT

```

tgggttcggga attcgaggct ctgacggagg agaatcggac gttgagggtg gccagtcctc 660
aatgtgtgga acaactggag aaacttcgaa tacagtatca gaagaggcag ggctcgtcct 720
aactttaaat ttttcagtgt gagcatacga ggctgatgac tgccctgtgc tggccaaaag 780
atTTTTatTTT taaatgaata gtgagtcaga tctattgctt ctctgtatta cccacatgac 840
aactgtctat aatgagttta ctgcttgcca gcttctagct tgagagaagg gatattttaa 900
atgagatcat taacgtgaaa ctattactag tataatgtttt tggagatcag aattcttttc 960
caaagatata tgTTTTTTTc ttttttagga agatatgac atgctgtaca acagggtaga 1020
aaatgataaa aatagactta ttgctgacct cag 1053

```

&lt;210&gt; 27

&lt;211&gt; 1212

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 2636759CB1

&lt;400&gt; 27

```

gacggcacgc aaggtcagcg tcacagatct gaccctaaaa ataggcctct gttgccagtc 60
gggggtggctg ggcgtgcggc tgctacatgc cccacggacc agaacctccc gacgcggcca 120
ggccccggca caccagctg cagaaaggag agaaaatccc ttggctctaa aatgacatct 180
ggagaagtga agacaagcct caagaatgcc tactcatctg ccaagaggct gtcgccgaag 240
atggaggagg aaggggagga ggaggactac tgcacccctg gagcctttga gctggagcgg 300
ctcttctgga agggcagtc ccagtacacc cacgtcaacg aggtctggcc caagctctac 360
attggcgatg aggcgacggc gctggaccgc tataggctgc agaaggcggg gttcacgcac 420
gtgctgaacg cggcccacgg ccgttggaac gtggacactg ggccccgact actaccgcga 480
catggacatc cagtaccacg gcgtggaggc ccgacgacct gccaccttt cgacctcagt 540
gtcttcttct acccggcggc agccttcate gacagagcgc taagcgacga ccacagtaag 600
atctctggtc actgcgtcat gggccgcagc cggtcagcca ccctggtoct ggcctacctg 660
atgatccaca aggacatgac cctggtggac gccatccagc aagtggccaa gaaccgctgc 720
gtcctccccg accggggctt tttgaagcag ctccgggagc tggacaagca gctggtgcag 780
cagaggcgac ggtcccagcg ccaggacggt gaggaggagg atgacagggg gctgtaggcc 840
cgactcacag ggccagcaga ggcacttggt gacagagggg agaggcagaa catagccctg 900
gcctaggact ccagagaagg gatggtgaaa ccgaagctcg actcttccaa accatcttgt 960
tcaacttccc catgtgtgct ggggacaggg aggaccaga gctgcccccg ggcagagctg 1020
agcgctcagc ctctcagcaa aatgggaggg acgggctccc cggctctggg tcacagagga 1080
gcatgccacg ctgcaccaag tcccctgctt tggTTTTgtt tttttggtga gaaggaagag 1140
ggaaaatttt ttaaattgtg gggcattatt tttgtaaata tccttcgggc tttgtttaaa 1200
aacaaaaaaa aa 1212

```

&lt;210&gt; 28

&lt;211&gt; 3093

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 2716815CB1

&lt;400&gt; 28

```

gcagagtgtc agcaggagcg cgagccagca agaggcgctt gcgcgatgtc cgggccccctg 60
agccgcggcg ctgagccagc cgggacggac atgcgcggga gggcgccgcg gggcagccgc 120
cgctcctccg ggggaatgaa agctactggt tgattttaaa gtgcctgggc ctcacagggt 180
tgagatgtc ccagaataag gcacaatgtc aatagcagga gttgctgtc aggagatcag 240
agtcccatTA aaaactggat ttctacataa tggccgagcc atggggaata tgaggaagac 300
ctactggagc agtcgcagt agtttaaaaa caacttttta aatattgacc cgataacat 360
ggcctacagt ctgaactctt ctgctcagga gcgcctaata ccacttgggc atgcttccaa 420
atctgctccg atgaatggcc actgctttgc agaaaatggt ccactcTcaa agtccagctt 480

```

PF-0683 PCT

```

gccccctctt cttattcccc caagtgaaaa cttgggacca catgaagagg atcaagttgt 540
atgtgggtttt aagaaactca cagtgaatgg ggtttgtgct tccacccttc cactgacacc 600
cataaaaaaac tccccctccc ttttccccctg tgccccctctt tgtgaacggg gttctaggcc 660
tcttccaccg ttgccaatct ctgaagccct ctctctggat gacacagact gtgaggtgga 720
attcctaact agctcagata cagacttcct tttagaagac tctacacttt ctgatttcaa 780
atatgatgtt cctggcaggc gaagcttccg tgggtgtgga caaatcaact atgcatattt 840
tgatacccca gctgtttctg cagcagatct cagctatgtg tctgacaaaa atggaggtgt 900
cccagatcca aatcctcctc cacctcagac ccaccgaaga ttaagaaggt ctcatcggg 960
accagctggc tcctttaaca agccagccat aaggatatcc aactgttgta tacacagagc 1020
ttctcctaac tccgatgaag acaaacctga gggtcccccc agagttccca tacctcctag 1080
accagtaaag ccagattata gaagatggtc agcagaagt acttcgagca cctatagtga 1140
tgaagacagg cctcccaaag taccgccaag agaacctttg tcaccgagta actcgcgcac 1200
accgagtccc aaaagccttc cgtcttacct caatggggtc atgccccga cacagagctt 1260
tgccccctgat cccaagtatg tcagcagcaa agcactgcaa agacagaaca gcgaaggatc 1320
tgccagtaag gttccttgca ttctgcccac tattgaaaat gggaagaagg ttagttcaac 1380
acattattac ctactacctg aacgaccacc atacctggac aaatatgaaa aatttttttag 1440
ggaagcagaa gaaacaaatg gaggcgcccc aatccagcca ttacctgctg actgcggtat 1500
actttcagcc acagaaaagc cagactcaaa aacaaaaatg gatctgggtg gccacgtgaa 1560
gcgtaaacat ttatcctatg tggtttctcc ttagaccttg gggtcatggt tcagcagagg 1620
ttacatagga gcaaatgggt ctcaattttc cagtttgatt gaagtgcaga gaaaaatccc 1680
ttagattgca aaataaaata gttgaactct ctgtcttcat gtggaagggt tagagcagtt 1740
gtgagatgct gttatgctga gaaaccctga ctttgttagt gttggaaaaa agtcttacaa 1800
gtctataatt taaagatgtg atggtgggga ggggaggatg gggaagcttt ttatatatgc 1860
atacattaca tacctatata taaacttgty gtataacct agaccatagc tgcagtttaa 1920
ccaattagtt actatcgtag agtaatatat attcagaata ataaactcaa gctggagaaa 1980
tgagtcctga tagactgaaa attgagcaaa tggaagaaga tacagtattg tttagatcag 2040
aatcattaaa aaatatTTTT gtttagtaag tttgaagatt tctggctttt aggccttttc 2100
tattttgttc catttatTTT tgcaggcaat cttttccatg gagggcaggg tatccattct 2160
ttaccatggg tgtacctgct taggttaaaa atcataccaa ggcctcatac ttccaggttt 2220
catgttgctt cttgttgagg gagggagagc aggttacttg gcaaccatat tgtcacctgt 2280
acctgtcaca catcttgaaa aataaaacga taatagaact agtgactaat tttcccttac 2340
agttcctgct tgggtcccacc cactgaagta gctcatcgta gtgcggggccg tattagaggc 2400
agtggggtac gttagactca gatggaaaag tattctaggt gccagtgtta ggatgtcagt 2460
tttacaaaat aatgaagcaa ttagctatgt gattgagagt tattgttttg ggatgtgtgt 2520
tgtggttttg cttttttttt ttagactgta ttaataaaca tacaacacaa gctggccttg 2580
tgttgctggg tcctattcag tatttcctgg ggattgtttg ctttttaagt aaaacacttc 2640
tgacctatag ctcatgtatg ctgaattcca gaggtcacat cagcatcttt ctgctttgaa 2700
aactctcaca gctgtggctg cttcacttag atgcagttag acacatagtt ggtgttccga 2760
ttttcacatc cttccatgta tttatcttga agagataagc acagaagaga aggtgtcac 2820
taacagaggt acattactgc aatgttctct taacagttaa acaagctgtt tacagtttaa 2880
actgctgaat attatttgag ctattttaaag cttattatat tttagtatga actaaatgaa 2940
ggttaaaaca tgcttaagaa aaatgcactg atttctgcat tatgtgtaca gtattggaca 3000
aaggatttta ttcattttgt tgcattattt tgaatattgt cttttcattt taataaagtt 3060
ataatactta tttatgataa aaaaaaaaaa aaa 3093

```